

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:13 ; Search time 244.117 Seconds
(Without alignments)
10530.153 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 eggsgaggtacgggggaattga.....cccagaataactgagag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------------------|-------------------|
| 1 | 1567.8 | 99.8 | 10818 | 4 US-09-949-016-13583 | Sequence 13583, A |
| 2 | 599 | 38.1 | 601 | 4 US-09-949-016-63416 | Sequence 63416, A |
| 3 | 482 | 30.7 | 735 | 3 US-09-602-877A-94 | Sequence 94, Appl |
| 4 | 482 | 30.7 | 1143 | 4 US-09-949-016-1841 | Sequence 1841, Ap |
| 5 | 482 | 30.7 | 1499 | 4 US-09-509-908-1 | Sequence 1, Appl1 |
| 6 | 480.6 | 30.6 | 601 | 4 US-09-949-016-63415 | Sequence 63415, A |
| 7 | 476 | 30.3 | 1476 | 2 US-08-824-874-2 | Sequence 2, Appl1 |
| 8 | 476 | 30.3 | 1476 | 2 US-09-210-084-2 | Sequence 2, Appl1 |
| 9 | 476 | 30.3 | 1476 | 4 US-09-764-762-2 | Sequence 2, Appl1 |
| 10 | 460 | 29.3 | 1504 | 3 US-09-280-116-1 | Sequence 1, Appl1 |
| 11 | 253 | 16.1 | 601 | 4 US-09-949-016-63414 | Sequence 63414, A |
| 12 | 223.4 | 14.2 | 68719 | 4 US-09-949-016-12799 | Sequence 12799, A |
| 13 | 223.4 | 14.2 | 68720 | 4 US-09-949-016-14296 | Sequence 14296, A |
| 14 | 222.8 | 14.2 | 32278 | 4 US-09-949-016-14575 | Sequence 14575, A |
| 15 | 222.6 | 14.2 | 176373 | 3 US-09-128-155-17 | Sequence 17, Appl |
| 16 | 222.2 | 14.1 | 601 | 4 US-09-949-016-66914 | Sequence 66914, A |
| 17 | 222.2 | 14.1 | 177668 | 4 US-09-949-016-13713 | Sequence 13713, A |
| 18 | 221.8 | 14.1 | 80858 | 4 US-09-949-016-12659 | Sequence 12659, A |
| 19 | 221.8 | 14.1 | 80859 | 4 US-09-949-016-15715 | Sequence 15715, A |
| 20 | 220.8 | 14.1 | 69701 | 4 US-09-949-016-14187 | Sequence 14187, A |
| 21 | 220.8 | 14.1 | 73308 | 4 US-09-949-016-16336 | Sequence 16336, A |
| 22 | 219.8 | 14.0 | 300402 | 4 US-09-949-016-13632 | Sequence 13632, A |
| 23 | 219.4 | 14.0 | 38009 | 4 US-09-949-016-13617 | Sequence 13617, A |
| 24 | 219.4 | 14.0 | 70000 | 3 US-09-851-896-3 | Sequence 3, Appl1 |
| 25 | 219.4 | 14.0 | 76399 | 4 US-09-949-016-16819 | Sequence 16819, A |
| 26 | 219.2 | 14.0 | 152331 | 3 US-09-128-155-16 | Sequence 16, Appl |
| 27 | 218.8 | 13.9 | 71574 | 4 US-09-949-016-15580 | Sequence 15580, A |

| | | | | | |
|----|-------|------|--------|-----------------------|-------------------|
| 28 | 218.6 | 13.9 | 25464 | 4 US-09-326-480A-4 | Sequence 4, Appl1 |
| 29 | 218.6 | 13.9 | 87039 | 4 US-09-949-016-15691 | Sequence 15691, A |
| 30 | 218.2 | 13.9 | 41755 | 4 US-09-949-016-15728 | Sequence 15728, A |
| 31 | 218 | 13.9 | 5695 | 4 US-09-949-016-13664 | Sequence 13664, A |
| 32 | 217.8 | 13.9 | 301828 | 4 US-09-949-016-13669 | Sequence 13669, A |
| 33 | 217.8 | 13.9 | 601 | 4 US-09-949-016-64323 | Sequence 64323, A |
| 34 | 217.8 | 13.9 | 36156 | 4 US-09-949-016-12128 | Sequence 12128, A |
| 35 | 217.8 | 13.9 | 36156 | 4 US-09-949-016-12361 | Sequence 13261, A |
| 36 | 217.8 | 13.9 | 46841 | 4 US-09-949-016-13466 | Sequence 13466, A |
| 37 | 217.6 | 13.9 | 246444 | 4 US-09-949-016-13113 | Sequence 13113, A |
| 38 | 217.6 | 13.9 | 321022 | 4 US-09-949-016-11852 | Sequence 11852, A |
| 39 | 217.6 | 13.9 | 321022 | 4 US-09-949-016-14166 | Sequence 14166, A |
| 40 | 217.2 | 13.8 | 101128 | 4 US-09-949-016-14293 | Sequence 14293, A |
| 41 | 216.8 | 13.8 | 10013 | 4 US-09-949-016-16474 | Sequence 16474, A |
| 42 | 216.6 | 13.8 | 44019 | 4 US-09-949-016-14802 | Sequence 14902, A |
| 43 | 216.4 | 13.8 | 461 | 3 US-09-404-879A-1 | Sequence 1, Appl1 |
| 44 | 216.4 | 13.8 | 461 | 3 US-09-404-879A-3 | Sequence 3, Appl1 |
| 45 | 216.4 | 13.8 | 461 | 4 US-09-338-933-1 | Sequence 1, Appl1 |

ALIGNMENTS

| | | | | | |
|---|------|---|------|--|--|
| RESULT 1 | | | | | |
| US-09-949-016-13583 | | | | | |
| Sequence 13583, Application US/09949016 | | | | | |
| Patent No. 6812335 | | | | | |
| GENERAL INFORMATION: | | | | | |
| APPLICANT: VENTNER, J. Craig et al. | | | | | |
| TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | | | |
| TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF | | | | | |
| FILE REFERENCE: CLO010307 | | | | | |
| CURRENT APPLICATION NUMBER: US/09/949,016 | | | | | |
| CURRENT FILING DATE: 2000-04-14 | | | | | |
| PRIOR APPLICATION NUMBER: 60/241,755 | | | | | |
| PRIOR FILING DATE: 2000-10-20 | | | | | |
| PRIOR APPLICATION NUMBER: 60/237,768 | | | | | |
| PRIOR FILING DATE: 2000-10-03 | | | | | |
| PRIOR APPLICATION NUMBER: 60/231,498 | | | | | |
| PRIOR FILING DATE: 2000-09-08 | | | | | |
| NUMBER OF SEQ ID NOS: 207012 | | | | | |
| SOFTWARE: FASTSEQ for Windows Version 4.0 | | | | | |
| SEQ ID NO 13583 | | | | | |
| LENGTH: 10818 | | | | | |
| TYPE: DNA | | | | | |
| ORGANISM: Human | | | | | |
| US-09-949-016-13583 | | | | | |
| Query Match | | | | | |
| Best Local Similarity 99.8%; Score 1567.8; DB 4; Length 10818; | | | | | |
| Matches 1569; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | | | | |
| QY | 1 | AGGAGAGTATGGGGAATTGAGACAGAGAAACAAATTAGTCCAGCGATGATTCTA | 60 | | |
| DB | 7243 | AGGAGAGTATGGGGAATTGAGACAGAGAAACAAATTAGTCCAGCGATGATTCTA | 7302 | | |
| QY | 61 | TTGGAGTGTATTTGCTCCCTAGAGACACTGGCAATACAGAGAGATTTTGGTTGTA | 120 | | |
| DB | 7243 | TTGGAGTGTATTTGCTCCCTAGAGACACTGGCAATACAGAGAGATTTTGGTTGTA | 7362 | | |
| QY | 121 | CAACTATATGAGGGGCACTTACTGCGACATATGATGATGCAAGTGGCTTCAAC | 180 | | |
| DB | 7303 | CAACTATATGAGGGGCACTTACTGCGACATATGATGATGCAAGTGGCTTCAAC | 7422 | | |
| QY | 181 | ATGCTATGATGACACAGGAGGCTTCAACAAACATTTATCAGATGCCAC | 240 | | |
| DB | 7423 | ATGCTATGATGACACAGGAGGCTTCAACAAACATTTATCAGATGCCAC | 7482 | | |
| QY | 241 | AGTCCCAATGACAGAACCCCTATCAGAGGCTTGAAACCGTATTTTGCAGAGGAG | 300 | | |
| DB | 7483 | AGTCCCAATGACAGAACCCCTATCAGAGGCTTGAAACCGTATTTTGCAGAGGAG | 7542 | | |
| QY | 301 | GTTAAGATGGGTGGTGGAGATGGGAGAGAGTGTGTCTCAGTAAGAAATTA | 360 | | |

[illegible]

| | | | |
|----|------|---|------|
| Db | 8623 | GGACTCAGGGCTCTGCTTCCCCCAATTTGGGCTGACCGGTCTCTCTAGTTGAACCTGGG | 8682 |
| Qy | 1441 | AACAATTTCCAAACTGTCCAGGGCGGGGTTGGTCTCAATCTCCCTGGGGCACTTTCA | 1500 |
| Db | 8683 | AACAATTTCCAAACTGTCCAGGGCGGGGTTGGTCTCAATCTCCCTGGGGCACTTTCA | 8742 |
| Qy | 1501 | TCTTCAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTATGTCGCAGAAT | 1560 |
| Db | 8743 | TCTTCAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTATGTCGCAGAAT | 8802 |
| Qy | 1561 | AAACTGAGAG 1571 | |
| Db | 8803 | AAACTGAGAG 8813 | |

```

RESULT 2
US-09-949-016-63416/C
; Sequence 63416, Application US/09949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63416
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63416

```

| Query Match | 38.1% | Score 599 | DB 4 | Length 601 |
|-----------------------|----------------|---|----------|------------|
| Best Local Similarity | 99.7% | Pred No. 3.4e-151 | | |
| Matches 599 | Conservative 1 | Mismatches 1 | Indels 0 | Gaps 0 |
| Qy | 482 | AGACCCCCCGAAATGATGTGTGACCAACAGAAATCTTGAAAGAGAAAGATGAGTGGAGA | 5411 | |
| Db | 601 | AGACCCCCCGAAATGATGTGTGACCAACAGAAATCTTGAAAGAGAAAGATGAGTGGAGA | 5422 | |
| Qy | 542 | GTTACAAATGGGCTCTAAAGCTTGAACCTTGAGGCAAGGCAATGGTGGCTACACGCTGTAA | 6010 | |
| Db | 541 | GTTACAAATGGGCTCTAAAGCTTGAACCTTGAGGCAAGGCAATGGTGGCTACACGCTGTAA | 4822 | |
| Qy | 602 | TCCCAACACTTTGGAGGCTGAGGTGGGCGAAATCACTTGAGGCCAGAGTTGAGACCAGC | 6615 | |
| Db | 481 | TCCCAACACTTTGGAGGCTGAGGTGGGCGAAATCACTTGAGGCCAGAGTTGAGACCAGC | 4222 | |
| Qy | 662 | CTGGCCCAATGCTGAAACCCCGCTCTTACAAAAAAATACAAAAAATTAGCCGGGTGTG | 7211 | |
| Db | 421 | CTGGCCCAATGCTGAAACCCCGCTCTTACAAAAAAATACAAAAAATTAGCCGGGTGTG | 3622 | |
| Qy | 722 | GTCATGAGCAACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCAAGAAATTGCTTGAAC | 7818 | |
| Db | 361 | GTCATGAGCAACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCAAGAAATTGCTTGAAC | 3022 | |
| Qy | 782 | CGGGAATGAGGCTGCACTGAGCTGAGGCTCAGGCCCATCTGGCTTCAAACCTGGGCAACAG | 8411 | |
| Db | 301 | YGGGAATGAGGCTGCACTGAGCTGAGGCTCAGGCCCATCTGGCTTCAAACCTGGGCAACAG | 2422 | |
| Qy | 842 | AGTAAAGCTCAATCTCAAAAAAAAGAGCTGAAATTGGAATGAAATATTAATPAACATTC | 9010 | |
| Db | 241 | AGTAAAGCTCAATCTCAAAAAAAAGAGCTGAAATTGGAATGAAATATTAATPAACATTC | 1822 | |
| Qy | 902 | TCCCTCTCTCCCTTTGGCTGCTGTCTCATCTCTGCTTTTTCGCAATTTCTTCATCTC | 9611 | |


```

Sequence 1, Application US/09509908
Patent No. 6589770
GENERAL INFORMATION:
APPLICANT: The Procter & Gamble Company, N/A N/A
TITLE OF INVENTION: A Protease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: T. David Reed
STREET: 5299 Spring Grove Avenue
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZIP: 45217-1087
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09509,908
FILING DATE: 28-Feb-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Reed, T David
REGISTRATION NUMBER: 32,931
REFERENCE/DOCKET NUMBER: AA-264F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 513-627-7025
TELEFAX: 513-627-6333
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 291..1172
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 489..1172
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..290
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1173..1499
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-509-908-1

```

| | Query Match | 30.7% | Score 482 | DB 4 | Length 1499 | |
|----|-----------------------|--|--------------------|----------|-------------|--|
| | Best Local Similarity | 100.0% | Pred. No. 1.5e-119 | | | |
| | Matches 482 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 | |
| Qy | 1090 | CAGGGTGAATTCGTGGGGGGCTGTGGTCTGCATGGCTCCCTGCAGGGACTGTGCTCGG | 1149 | | | |
| Db | 1014 | CAGGGTGAATTCGTGGGGGGCTGTGGTCTGCATGGCTCCCTGCAGGGACTGTGCTCGG | 1073 | | | |
| Qy | 1150 | GGAAGTTTACCTTTGTGCCCCGGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC | 1209 | | | |
| Db | 1074 | GGAAGTTTACCTTTGTGCCCCGGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC | 1133 | | | |
| Qy | 1210 | ACCAAGTGAATCCAGAAACCATCCAGAGCAACTCTGAGCATCTCCAGAGCTCAGACCA | 1269 | | | |
| Db | 1134 | ACCAAGTGAATCCAGAAACCATCCAGAGCAACTCTGAGCATCTCCAGAGCTCAGACCA | 1193 | | | |
| Qy | 1270 | CCGGCATCCCAACCTGCTGCAGGGACAGCCCTGACACTCTTTGAGACCTCATTTCTCTTC | 1329 | | | |
| Db | 1194 | CCGGCATCCCAACCTGCTGCAGGGACAGCCCTGACACTCTTTGAGACCTCATTTCTCTTC | 1253 | | | |
| Qy | 1330 | CCAGAGATTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCCTCTGGAATCAGG | 1389 | | | |

| | | | |
|----|------|---|------|
| Db | 1255 | CCAGAGATGTTGAGAAATGTTCAATCTCTCGACAGCCCTGACCCCAAGTCTCCTGGACTCAGG | 1313 |
| OY | 1390 | GTCTGCTTCCCCCAATTTGGGCTGACCGGTGTCTCTCAATTGAAACCTTGGGAAACAATTTTC | 1449 |
| Db | 1314 | GTCGTCTTCCCCCAATTTGGGCTGACCGGTGTCTCTCAATTGAAACCTTGGGAAACAATTTTC | 1373 |
| OY | 1450 | CAAAACTGTCCAGGGCGGGGGTTGCGTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC | 1509 |
| Db | 1374 | CAAAACTGTCTCAGGGCGGGGGTTGCGTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC | 1433 |
| OY | 1510 | TCAGGGCCCATCCCTTCTCTGACAGCTTGAACCCAAATTTAGTCCCAAGAAATTAACCTGAGA | 1569 |
| Db | 1434 | TCAGGGCCCATCCCTTCTCTGACAGCTTGAACCCAAATTTAGTCCCAAGAAATTAACCTGAGA | 1493 |
| OY | 1570 | AG 1571 | |
| Db | 1494 | AG 1495 | |

```

RESULT 6
US-09-949-016-63415/c
: Sequence 63415, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 63415
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-63415

```

| Query Match | 30.6% | Score 480.6 | DB 4 | Length 601 |
|-----------------------|----------------|---|----------|------------|
| Best Local Similarity | 99.8% | Pred. No. 2.3e-119 | | |
| Matches 480 | Conservative 1 | Mismatches 0 | Indels 0 | Gaps 0 |
| Qy | 1 | AGGAGGGATGGGGGAAATTGAAAGACAGGAAACAAATTATAGTCCAGCGAATGGATTTCTA | 60 | |
| Db | 481 | AGGAGGGATGGGGGAAATTGAAAGACAGGAAACAAATTATAGTCCAGCGAATGGATTTCTA | 422 | |
| Qy | 61 | TTGGGAGGATTTCTGCCCCCTAGAAAGACACTGGCAATACAGAGACATTTTGGTTGTCA | 120 | |
| Db | 421 | TTGGGAGGATTTCTGCCCCCTAGAAAGACACTGGCAATACAGAGACATTTTGGTTGTCA | 362 | |
| Qy | 121 | CAACTATATGGAGGGGCATTACTGCGCACTAATGATAGATGCCAAGTGTCTGTTCAAC | 180 | |
| Db | 361 | CAACTATATGGAGGGGCATTACTGCGCACTAATGATAGATGCCAAGTGTCTGTTCAAC | 302 | |
| Qy | 181 | ATGCTATGATGCACACGCGCAGGCTTCCACAAACCAATTATCCAGCTTCAGATGCCAC | 240 | |
| Db | 301 | RTGCTATGATGCACACGCGCAGGCTTCCACAAACCAATTATCCAGCTTCAGATGCCAC | 242 | |
| Qy | 241 | AGTCCCGCAGATCGAGAAACCTCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG | 300 | |
| Db | 241 | AGTCCCGCAGATCGAGAAACCTCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG | 182 | |
| Qy | 301 | GTATTAAGATGGGTTTGTGAGAAATGGGAAAGAAAGTGTGTGTCCAGTAAAGAAATTA | 360 | |
| Db | 181 | GTATTAAGATGGGTTTGTGAGAAATGGGAAAGAAAGTGTGTGTCCAGTAAAGAAATTA | 122 | |
| Qy | 361 | GGCTTCGACAGCGCTGGAGGGGAGAGTGAAGAGAAAAGGAGCGGAGAGATTCACGATTA | 420 | |

Db 121 GGCCTGCACAGGCTGGAGGGGAGAGTGAAGAGAAAGGAGCGGAGAGATACACATGA 62
Qy 421 GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAATTCAGATGTGGAGAGAAAGGCTCA 480
Db 61 GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAATTCAGATGTGGAGAGAAAGGCTCA 2
Qy 481 C 481
Db 1 C 1

RESULT 7
US-08-824-874-2
Sequence 2, Application US/08824874
Patent No. 5962300
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-08-824-874-2

Query Match 30.3%; Score 476; DB 2; Length 1476;
Best Local Similarity 100.0%; Pred. No. 66-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCCTGACGGGAGCTGTCTGG 1149
Db 1001 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCCTGACGGGAGCTGTCTGG 1060
Qy 1150 GGAATTAACCTTGTGCGCGGACCAAGACGCGGGTGTCAACAGAACTCTGCAATTC 1209
Db 1061 GGAGATTAACCTTGTGCGCGGACCAAGACGCGGGTGTCAACAGAACTCTGCAATTC 1120
Qy 1210 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGCTCAGACA 1269
Db 1121 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGCTCAGACA 1180

Qy 1270 CCGGCAATCCCACTGCTGACAGGAGACGCCCTGACACTCCTTTCAGACCTCAATTCCTTC 1329
Db 1181 CCGGCAATCCCACTGCTGACAGGAGACGCCCTGACACTCCTTTCAGACCTCAATTCCTTC 1240
Qy 1330 CCAGAGATGTGGAAATGTTCACTCTCCAGCCCTGAGCCCAATGTTCTCTGAGACTCAGG 1389
Db 1241 CCAGAGATGTGGAAATGTTCACTCTCCAGCCCTGAGCCCAATGTTCTCTGAGACTCAGG 1300
Qy 1390 GTCTGCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAAACAATTC 1449
Db 1301 GTCTGCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAAACAATTC 1360
Qy 1450 CAAAACGTCCAGAGGGGGGGGCTTGCCTTCAATCCCTGGGGACATTCATCTCAAGC 1509
Db 1361 CAAAACGTCCAGAGGGGGGGGCTTGCCTTCAATCCCTGGGGACATTCATCTCAAGC 1420
Qy 1510 TCAGGGCCCATCCCTTCTGTGACGCTTGACCCAAATTTAGTCCAGAAATAACT 1565
Db 1421 TCAGGGCCCATCCCTTCTGTGACGCTTGACCCAAATTTAGTCCAGAAATAACT 1476

RESULT 8
US-09-210-084-2
Sequence 2, Application US/09210084
Patent No. 6197511
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-09-210-084-2

Query Match 30.3%; Score 476; DB 3; Length 1476;
Best Local Similarity 100.0%; Pred. No. 66-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCCTGACGGGAGCTGTCTGG 1149

DB 1001 CAGGGGATGATTCGGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTGTCTGG 1060
QY 1150 GGAGATTACCTCTGTGCTCCGGGGCCCAAGACGCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 1061 GGAGATTACCTCTGTGCTCCGGGGCCCAAGACGCGGGGTGTCTACAGAACTCTGCAAGTTC 1120
QY 1210 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGAGACTCAGACACA 1269
DB 1121 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGAGACTCAGACACA 1180
QY 1270 CCGGCAATCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTTCAAGACCTTCAATTCCTTC 1329
DB 1181 CCGGCAATCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTTCAAGACCTTCAATTCCTTC 1240
QY 1330 CCAGAGATGTTGAGAAATGTTCACTCTGACAGCCCGCCGACCCCAATGTCTCCTGAGACTCAGG 1389
DB 1241 CCAGAGATGTTGAGAAATGTTCACTCTGACAGCCCGCCGACCCCAATGTCTCCTGAGACTCAGG 1300
QY 1390 GTCTGCTTCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTGATGTAAACCTGGAACAATTC 1449
DB 1301 GTCTGCTTCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTGATGTAAACCTGGAACAATTC 1360
QY 1450 CAAAATGTCAGGCGGGGGGTGCTTCATCTCTCTGAGGCACTTTCATCTCAAGC 1509
DB 1361 CAAAATGTCAGGCGGGGGGTGCTTCATCTCTCTGAGGCACTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1565
DB 1421 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1476

RESULT 9

US-09-764-762-2

; Sequence 2, Application US/09764762

; Patent No. 6472195

; GENERAL INFORMATION:

; APPLICANT: Hallman, Jennifer L.

; La1, Preeti

; TITLE OF INVENTION: NOVEL KALLIKREIN

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/764,762

; FILING DATE: 16-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/210,084

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0252 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1476 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 820694
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Query Match 30.3%; Score 476; DB 4; Length 1476;
Best Local Similarity 100.0%; Pred. No. 6e-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTGTCTGG 1149
DB 1001 CAGGGTGAATTCGGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTGTCTGG 1060
QY 1150 GGAGATTACCTCTGTGCTCCGGGGCCCAAGACGCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 1061 GGAGATTACCTCTGTGCTCCGGGGCCCAAGACGCGGGGTGTCTACAGAACTCTGCAAGTTC 1120
QY 1210 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGAGACTCAGACACA 1269
DB 1121 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGAGACTCAGACACA 1180
QY 1270 CCGGCAATCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTTCAAGACCTTCAATTCCTTC 1329
DB 1181 CCGGCAATCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTTCAAGACCTTCAATTCCTTC 1240
QY 1330 CCAGAGATGTTGAGAAATGTTCACTCTGACAGCCCGCCGACCCCAATGTCTCCTGAGACTCAGG 1389
DB 1241 CCAGAGATGTTGAGAAATGTTCACTCTGACAGCCCGCCGACCCCAATGTCTCCTGAGACTCAGG 1300
QY 1390 GTCTGCTTCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTGATGTAAACCTGGAACAATTC 1449
DB 1301 GTCTGCTTCCCACTGTGCTGAGGGAAGAGCCCTGACACTCTCTGATGTAAACCTGGAACAATTC 1360
QY 1450 CAAAATGTCAGGCGGGGGGTGCTTCATCTCTCTGAGGCACTTTCATCTCAAGC 1509
DB 1361 CAAAATGTCAGGCGGGGGGTGCTTCATCTCTCTGAGGCACTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1565
DB 1421 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1476

RESULT 10

US-09-280-116-1

; Sequence 1, Application US/09280116A

; Patent No. 6331427

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

; FILE REFERENCE: 5800-24, 035800/176965

; CURRENT APPLICATION NUMBER: US/09/280,116A

; CURRENT FILING DATE: 1999-03-26

; NUMBER OF SEQ ID NOS: 268

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1504

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: trypsin-like serine proteases

US-09-280-116-1

Query Match 29.3%; Score 460; DB 3; Length 1504;
Best Local Similarity 99.6%; Pred. No. 1.2e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1090 CAGGGTGAATTCGGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTGTCTGG 1148
DB 1015 CAGGGTGAATTCGGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTGTCTGG 1074
QY 1149 GGAGATTACCTCTGTGCTCCGGGGCCCAAGACGCGGGGTGTCTACAGAACTCTGCAAGTT 1208

Db 1075 GGGAGATTACCTTGTGTCGCCGCCAAGACCGGGTGTCTACAGAACTCTGCAAGTT 1134
1209 CAGCAAGTGTATCCAGAAACATCCAGGCCAACTCTGAGTATCTCCAGACTCAGAC 1268
1135 CACCAAGTGTATCCAGAAACATCCAGGCCAACTCTGAGTATCTCCAGACTCAGAC 1194
1269 ACAGGCACTCCCACTCTGAGGAGCAGGCTGACACTCTTCAAGACCTCATTGCTT 1128
1195 ACCGGCATCCCACTCTGAGGAGCAGGCTGACACTCTTCAAGACCTCATTGCTT 1254
1329 CCGAGAGATGTGAATGATGTTCTCTCCAGCCCTGACCCCAATGCTCTGAGTCTAG 1388
1255 CCGAGAGATGTGAATGATGTTCTCTCCAGCCCTGACCCCAATGCTCTGAGTCTAG 1314
1389 GGTCTGCTTCCCACTCTGAGGAGCAGGCTGACACTCTTCAAGACCTCATTGCTT 1448
1315 GGTCTGCTTCCCACTCTGAGGAGCAGGCTGACACTCTTCAAGACCTCATTGCTT 1374
1449 CCAAACTGTCCAGGGGGGGGTTGCTCTCAATCTCCCTGGGGCACTTTCATCTGCAAG 1508
1375 CCAAACTGTCCAGGGGGGGGTTGCTCTCAATCTCCCTGGGGCACTTTCATCTGCAAG 1434
1509 CTAGGGCCCATCTCTCTGAGCTCTGACCCCAATTTAGT-CCCAAGAAATTAAGTGA 1567
1435 CTAGGGCCCATCTCTCTGAGCTCTGACCCCAATTTAGTCCCAAGAAATTAAGTGA 1494
Qy 1568 GAAG 1571
Db 1495 GAAG 1498

RESULT 11

US-09-949-016-63414/c
; Sequence 63414, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63414
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63414

Query Match 16.1%; Score 253; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 3, 5e-58;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAGGTATGGGAATTGAAGACAGAAACAAATTTCAAGCAAGCAATTTCTA 60
Db 253 AGGAGGTATGGGAATTGAAGACAGAAACAAATTTCAAGCAAGCAATTTCTA 194
Qy 61 TTGGAGTGTATTTGCCCCCTAGAAAGACTGCAATTCAGAGACATTTTGGTGTCA 120
Db 193 TTGGAGTGTATTTGCCCCCTAGAAAGACTGCAATTCAGAGACATTTTGGTGTCA 134
Qy 121 CAACTATATGAGGGGCACTTACTGCGCACTAATGATGATGCCAGTGTCTGTTCAAC 180
Db 133 CAACTATATGAGGGGCACTTACTGCGCACTAATGATGATGCCAGTGTCTGTTCAAC 74
Qy 181 ATGCTATGATGACAGCGGAGGCTCCAGCAAAACATTTATCAGTTGAGTGGCCAC 240

Db 73 ATGCTATGATGACAGCGGAGGCTCCAGCAAAACATTTATCAGTTGAGTGGCCAC 14
Qy 241 AGTGGCCAGATCG 253
Db 13 AGTGGCCAGATCG 1

RESULT 12

US-09-949-016-12799
; Sequence 12799, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12799
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12799

Query Match 14.2%; Score 223.4; DB 4; Length 68719;
Best Local Similarity 86.0%; Pred. No. 2, 8e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 570 TGGAGGCCAGGCACTGTGCTCAGCGCTGTATTCCTCAACATTT-AGAGGCTGAGGTGG 628
Db 57314 TGGAGGCTGGGCACTGTGCTCAGCGCTGTATTCCTCAACATTTGGAGGCCAGG 57373
Qy 629 CGAATCACTTGAAGCCAGGAGTTGAGACAGGCTGCGCAACATGAGTGAACCCGCTTC 688
Db 57374 TGGATCACTTGAAGTTGAGAGTTCAAGACAGGCTGCGCAACATGAGTGAACCCGCTTC 57433
Qy 689 TACAAAAAATACAAAAAATTAAGCGGCTGTGTGATGAGACCTGTAGTCAACAGTAC 748
Db 57434 TACTAAAAAATACAAAAAATTAAGCGGCTGTGTGATGAGACCTGTAGTCAACAGTAC 57493
Qy 749 TTGGAGGCTGAGGAGGAGAAATTTGTAACCCGGGAGATGAGGCTGCAGTGAGCTGA 808
Db 57494 TTGGAGGCTGAGGAGGAGAAATTTGTAACCTGGAGACAGAGGTTGCAGTGAGCTGA 57553
Qy 809 GGTCAAGCCACTGCGCTCCCACTGGGCAAGAGTGAAGTCAATCTCAAAAAA 868
Db 57554 GATCGTCCACTGTACTTCACTGGGCAAGAGGAGGAGTCTCTTCAAAAAA 57613
Qy 869 A 869
Db 57614 A 57614

RESULT 13

US-09-949-016-14296
; Sequence 14296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14296
LENGTH: 68720
TYPE: DNA
ORGANISM: Human
US-09-949-016-14296

Query Match 14.2%; Score 223.4; DB 4; Length 68720;
Best Local Similarity 86.0%; Pred. No. 2.8e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 570 TGGAGGCCAGGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GAGGCTGAGGTGG 628
DB 57314 TGGAGGCTGGGCACGGTGGCTCAAGCTGTATCCAGACACTTTGGGAGGCCAAGGCAAG 57373
QY 629 CGAATCACTTGAGGCCAGAGATTGAGACCAAGCTTGCCCAACATGTGTAACCCCGTCTC 688
DB 57374 TGGATCACTGAGGTCAAGAGTTCAAGACCAAGCTTGCCCAACATGTGCGAACCCTGCTC 57433
QY 689 TACAAAAAATACAAAAATTTAGCCGGGTGTGATGGAACAACCTGTGTCAAGCTAC 748
DB 57434 TACTAAAAAATACAAAAATTTAGCCAGCATGTGTGTGGACCTGTATATCCCAAGTGA 57493
QY 749 TTGGAGGCTGAGCAGAGAAATTTGTTGAACCCGGAGATGAGGCTGAGTGAAGCTGA 808
DB 57494 TTGGAGGCTGAGCAGAGAAATTTGTTGAACCTGTGAACCTGGAGACAGAGTTGAGTGA 57553
QY 809 GGTGAGGCCACTCGCTCCCACTGGGCAACAGATGAAGTCTCAATCTCAAAAAA 868
DB 57554 GATCGGCGCACTGTACTCCACCTGGGCAACAGAGAGGACTCCCTCTCAGAAAAA 57613
QY 869 A 869
DB 57614 A 57614

RESULT 14
US-09-949-016-14575
Sequence 14575, Application US/09949016
Patent No. 6812338
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14575
LENGTH: 32278
TYPE: DNA
ORGANISM: Human
US-09-949-016-14575

Query Match 14.2%; Score 222.8; DB 4; Length 32278;
Best Local Similarity 85.8%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 42; Indels 1; Gaps 1;
QY 569 TTGGAGGCCAGGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GAGGCTGAGGTGA 627

DB 13360 TTGTTGGCAGGCGTAGTGGCTCAGCGCTGTATCCAGACACTTTGGGAGACCAAGTGG 13419
QY 628 GGAATCACTTGAAGCCAGGAATTTGAGACCAAGCTTGCCCAACATGTGTAACCCCGTCT 687
DB 13420 GGAATCACTTGAAGCCAGGAATTTGAGACCAAGCTTGCCCAACATGTGTAACCCCGTCT 13479
QY 688 CTACAAAAAATTAACAAAAATTTAGCCGGGTGTGTGATGACACTGTAGTCAAGACTA 747
DB 13480 CTACAAAAAATTAACAAAAATTTAGCCAGGTGTGTGAGAGGTGCTGTATGCTCCAGCTA 13539
QY 748 CTTGGAGGCTGAGGAGAGAAATTTGTAACCCGGAGATGAGAGCTGCACTGAGCTG 807
DB 13540 CTTGGAGGCTGAGGAGAGAAATTTGTAACCCGGAGATGAGAGCTGCACTGAGCTG 13539
QY 808 AGTCAAGCCACTGCGCTCCCACTGGGCAACAGATGAGACTCATCTCAAAAAA 867
DB 13600 AGATGCGCCACTGTCACTCCAGCTGGGCAACAGAGCTGCTCAAAACAAAC 13659
QY 868 AA 869
DB 13660 AA 13661

RESULT 15
US-09-128-155-17
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 14.2%; Score 222.6; DB 3; Length 176373;
Best Local Similarity 84.0%; Pred. No. 7.2e-49;
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 558 AAGGTTGAATTTGAGGCGCAGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GGA 616
DB 127012 AAGGTTGAATTTGAGGCGCAGCATGTGTGCTCAGCGCTGTATCCCAACACTTTGGGA 127071
QY 617 GGTGAGGTGGGCAATCACTTGAGGCCAGAGTTGAGACCAAGCTTGCCCAACATGTGTG 676
DB 127072 GGTGAGGTGGGCAATCACTTGAGGCCAGAGTTGAGACCAAGCTTGCCCAACATGTGTG 127131
QY 677 AAACCCCGTCTTACAAAAAATTAACAAAAATTTAGCCGGGTGTGTGATGGAACACTGT 736
DB 127132 AAACCCCGTCTTACAAAAAATTAACAAAAATTTAGCCGGGTGTGTGATGGAACACTGT 127191
QY 737 AGTCAAGCTACTTGAAGGCTGAGGAGAGATTTGCTTGAACCCGGAGATGAGGCT 796
DB 127192 AATCCAGCTACAGGAGGCTGAGGAGAGATTCCTTGAACCCGGAGAGAGATTT 127251
QY 797 GAGTGAAGTGAAGTCAAGGCCACTGCTCCAACTGGGCAACAGATGAAGCTCACTCACT 856

Db 127252 GCAGTGAGCCAGTTCAAGCCACTGCACCCGAGCCTGGGCAACAGAGCAAGACTTTGTCT 127311
Qy 857 CAAAAAAAAAAAA 869
Db 127312 CAAAAAAAAAAAA 127324

Search completed: February 27, 2005, 05:14:39
Job time : 250.117 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:12 ; Search time 845.613 Seconds
(Without alignments)
11008.461 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 agggaggtatgggggaattga.....ccagaataactgaggaag 1571

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 541 | 34.4 | 567 | 16 | US-10-029-386-7604 Sequence 7604, Ap |
| 2 | 482 | 30.7 | 586 | 9 | US-09-954-531-849 Sequence 849, App |
| 3 | 482 | 30.7 | 586 | 9 | US-09-954-531-1266 Sequence 1266, Ap |
| 4 | 482 | 30.7 | 735 | 9 | US-09-745-288-94 Sequence 94, Appl |
| 5 | 482 | 30.7 | 735 | 17 | US-10-453-919-94 Sequence 94, Appl |
| 6 | 482 | 30.7 | 1260 | 17 | US-10-172-118-1411 Sequence 1411, Ap |
| 7 | 482 | 30.7 | 1260 | 17 | US-10-295-027-515 Sequence 515, Appl |
| 8 | 482 | 30.7 | 1260 | 17 | US-10-173-999-53 Sequence 53, Appl |
| 9 | 482 | 30.7 | 1260 | 17 | US-10-342-887-1411 Sequence 1411, Ap |
| 10 | 482 | 30.7 | 1381 | 9 | US-09-739-907-52 Sequence 52, Appl |
| 11 | 482 | 30.7 | 1381 | 11 | US-09-938-671-52 Sequence 52, Appl |

ALIGNMENTS

| | | | | | | |
|----|-----|------|------|----|--------------------|-------------------|
| 12 | 482 | 30.7 | 1439 | 9 | US-09-739-907-53 | Sequence 53, Appl |
| 13 | 482 | 30.7 | 1439 | 11 | US-09-938-671-53 | Sequence 53, Appl |
| 14 | 482 | 30.7 | 1481 | 18 | US-10-723-860-5032 | Sequence 5032, Ap |
| 15 | 482 | 30.7 | 1516 | 9 | US-09-739-907-37 | Sequence 37, Appl |
| 16 | 482 | 30.7 | 1516 | 11 | US-09-938-671-37 | Sequence 37, Appl |
| 17 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 18 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 19 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 20 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 21 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 22 | 482 | 30.7 | 1570 | 9 | US-09-989-722-308 | Sequence 308, App |
| 23 | 482 | 30.7 | 1570 | 9 | US-09-991-073-308 | Sequence 308, App |
| 24 | 482 | 30.7 | 1570 | 9 | US-09-990-442-308 | Sequence 308, App |
| 25 | 482 | 30.7 | 1570 | 9 | US-09-991-163-308 | Sequence 308, App |
| 26 | 482 | 30.7 | 1570 | 9 | US-09-993-604-308 | Sequence 308, App |
| 27 | 482 | 30.7 | 1570 | 9 | US-09-990-456-308 | Sequence 308, App |
| 28 | 482 | 30.7 | 1570 | 9 | US-09-989-721-308 | Sequence 308, App |
| 29 | 482 | 30.7 | 1570 | 9 | US-09-992-598-308 | Sequence 308, App |
| 30 | 482 | 30.7 | 1570 | 9 | US-09-989-723-308 | Sequence 308, App |
| 31 | 482 | 30.7 | 1570 | 9 | US-09-989-735-308 | Sequence 308, App |
| 32 | 482 | 30.7 | 1570 | 9 | US-09-990-444-308 | Sequence 308, App |
| 33 | 482 | 30.7 | 1570 | 9 | US-09-991-181-308 | Sequence 308, App |
| 34 | 482 | 30.7 | 1570 | 9 | US-09-989-730-308 | Sequence 308, App |
| 35 | 482 | 30.7 | 1570 | 9 | US-09-990-436-308 | Sequence 308, App |
| 36 | 482 | 30.7 | 1570 | 9 | US-09-993-687-308 | Sequence 308, App |
| 37 | 482 | 30.7 | 1570 | 10 | US-09-989-724-308 | Sequence 308, App |
| 38 | 482 | 30.7 | 1570 | 10 | US-09-997-653-308 | Sequence 308, App |
| 39 | 482 | 30.7 | 1570 | 10 | US-09-989-724-308 | Sequence 308, App |
| 40 | 482 | 30.7 | 1570 | 10 | US-09-989-728-308 | Sequence 308, App |
| 41 | 482 | 30.7 | 1570 | 10 | US-09-990-441-308 | Sequence 308, App |
| 42 | 482 | 30.7 | 1570 | 10 | US-09-993-667-308 | Sequence 308, App |
| 43 | 482 | 30.7 | 1570 | 10 | US-09-997-428-308 | Sequence 308, App |
| 44 | 482 | 30.7 | 1570 | 10 | US-09-997-666-308 | Sequence 308, App |
| 45 | 482 | 30.7 | 1570 | 10 | US-09-990-438-308 | Sequence 308, App |

RESULT 1
US-10-029-386-7604
; Sequence 7604, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7604
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: NT HIT: AP135028.1, EVALUATE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: W73140.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUATE 6.00e-27
; US-10-029-386-7604
Query Match 34.4%; Score 541; DB 16; Length 567;
Best Local Similarity 100.0%; Pred. No. 7.2e-150;

LOCATION: (1) ... (586)
OTHER INFORMATION: n=a,t,g or c
US-09-954-531-1266

Query Match 30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2,5e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 1149
DB 501 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 442
QY 1150 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 441 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 1269
DB 381 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 322
QY 1270 CCGGATCCCACTCTGCTGACAGGAGACGCTTGAACCTCTTTCAGACCTTCAATCTTC 1329
DB 321 CCGGATCCCACTCTGCTGACAGGAGACGCTTGAACCTCTTTCAGACCTTCAATCTTC 262
QY 1330 CCAGAGATGTTGAATGTTCAATCTCTCAGGCCCTGACCCCATGTCCTGAGACTGAG 1389
DB 261 CCAGAGATGTTGAATGTTCAATCTCTCAGGCCCTGACCCCATGTCCTGAGACTGAG 202
QY 1390 GTCTGCTTCCCACTATGAGGCTGACCGGTCTCTCTAGTTGAACCTTGGGAACTATTC 1449
DB 201 GTCTGCTTCCCACTATGAGGCTGACCGGTCTCTCTAGTTGAACCTTGGGAACTATTC 142
QY 1450 CAAAATCTGTCAGGGGCGGGGTTGCTCTCAATCTCCCTGGGACATTTTCACTCAAGC 1509
DB 141 CAAAATCTGTCAGGGGCGGGGTTGCTCTCAATCTCCCTGGGACATTTTCACTCAAGC 82
QY 1510 TCAGGGCCCATCCCTTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 1569
DB 81 TCAGGGCCCATCCCTTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 4

US-09-745-288-94/c
Sequence 94, Application US/09745288
Patent No. US20010018058A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.446D1
CURRENT APPLICATION NUMBER: US/09/745,288
CURRENT FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 101
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 94
LENGTH: 735
TYPE: DNA
ORGANISM: Homo sapien
US-09-745-288-94

Query Match 30.7%; Score 482; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 1149
DB 500 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 441

QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 381
QY 1210 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 1269
DB 380 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 321
QY 1270 CCGGATCCCACTCTGCTGACAGGAGACGCTTGAACCTCTTTCAGACCTTCAATCTTC 1329
DB 320 CCGGATCCCACTCTGCTGACAGGAGACGCTTGAACCTCTTTCAGACCTTCAATCTTC 261
QY 1330 CCAGAGATGTTGAATGTTCAATCTCTCAGGCCCTGACCCCATGTCCTGAGACTGAG 1389
DB 260 CCAGAGATGTTGAATGTTCAATCTCTCAGGCCCTGACCCCATGTCCTGAGACTGAG 201
QY 1390 GTCTGCTTCCCACTATGAGGCTGACCGGTCTCTCTAGTTGAACCTTGGGAACTATTC 1449
DB 200 GTCTGCTTCCCACTATGAGGCTGACCGGTCTCTCTAGTTGAACCTTGGGAACTATTC 141
QY 1450 CAAAATCTGTCAGGGGCGGGGTTGCTCTCAATCTCCCTGGGACATTTTCACTCAAGC 1509
DB 140 CAAAATCTGTCAGGGGCGGGGTTGCTCTCAATCTCCCTGGGACATTTTCACTCAAGC 81
QY 1510 TCAGGGCCCATCCCTTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 1569
DB 80 TCAGGGCCCATCCCTTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 21
QY 1570 AG 1571
DB 20 AG 19

RESULT 5

US-10-453-919-94/c
Sequence 94, Application US/10453919
Publication No. US20040033230A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.446C7
CURRENT APPLICATION NUMBER: US/10/453,919
CURRENT FILING DATE: 2003-06-03
NUMBER OF SEQ ID NOS: 121
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 94
LENGTH: 735
TYPE: DNA
ORGANISM: Homo sapien
US-10-453-919-94

Query Match 30.7%; Score 482; DB 17; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 1149
DB 500 CAGGATATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGACAGGACTGTCTCTGG 441
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 381
QY 1210 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 1269
DB 380 ACCAATGATTCAGAGAAACCATCCAGGCCAATCTCTGAGTCATCCAGAGACTCAGACA 321

QY 1270 CCGGCAATCCCACTGCTGAGGAGACCCCTGACACTCTCTTCAAGACCTCTTC 1329
DB 320 CCGGCAATCCCACTGCTGAGGAGACCCCTGACACTCTCTTCAAGACCTCTTC 261
QY 1330 CCGAAGATGTTGAGATGTTCACTCTCCAGCCCGCCGACCCCATGCTCTCTGAGCTCAGG 1389
DB 260 CCGAAGATGTTGAGATGTTCACTCTCCAGCCCGCCGACCCCATGCTCTCTGAGCTCAGG 201
QY 1390 GTCGTCTTCCCACTGAGGCTGACCGGCTCTCTCTAGTTGAACCTTGGAACAATTC 1449
DB 200 GTCGTCTTCCCACTGAGGCTGACCGGCTCTCTCTAGTTGAACCTTGGAACAATTC 141
QY 1450 CAAACTGTCAGAGGCGGGGCTTGCCTCTCAATCTCCGAGGACCTTTCATCTCAAGC 1509
DB 140 CAAACTGTCAGAGGCGGGGCTTGCCTCTCAATCTCCGAGGACCTTTCATCTCAAGC 81
QY 1510 TCAGGCGCCATCCCTCTCTGAGCTCTGACCCCAATTAAGTCCCGAATAAATGAGA 1569
DB 80 TCAGGCGCCATCCCTCTCTGAGCTCTGACCCCAATTAAGTCCCGAATAAATGAGA 21
QY 1570 AG 1571
DB 20 AG 19

RESULT 6

US-10-172-118-1411
; Sequence 1411, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1411
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; DATABASE ACCESSION NUMBER: NM_012427
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.4e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATGATTCGGGGGAGCTGAGTCTGCAATGCTCTCCCTGAGGAGACTCGTCTTCG 1149
DB 766 CAGGATGATTCGGGGGAGCTGAGTCTGCAATGCTCTCCCTGAGGAGACTCGTCTTCG 825
QY 1150 GGAAGATTACCTTGTGCTCCGAGCCCAAGAGCCGAGTGTCTAGACGAACCTCTGAGTTC 1209
DB 826 GGAAGATTACCTTGTGCTCCGAGCCCAAGAGCCGAGTGTCTAGACGAACCTCTGAGTTC 885
QY 1210 ACCAAGTGTATCCAGAAACCAATCCAGGCAACTCTCTGATGATCCAGAGTCAAGACA 1269
DB 886 ACCAAGTGTATCCAGAAACCAATCCAGGCAACTCTCTGATGATCCAGAGTCAAGACA 945
QY 1270 CCGGATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTCAGACCTCTATTCCTTC 1329
DB 946 CCGGATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTCAGACCTCTATTCCTTC 1005

QY 1330 CCGAAGATGTTGAGATGTTCACTCTCCAGCCCGCCGACCCCATGCTCTCTGAGCTCAGG 1389
DB 1006 CCGAAGATGTTGAGATGTTCACTCTCCAGCCCGCCGACCCCATGCTCTCTGAGCTCAGG 1065
QY 1390 GTCGTCTTCCCACTGAGGCTGACCGGCTCTCTCTAGTTGAACCTTGGAACAATTC 1449
DB 1066 GTCGTCTTCCCACTGAGGCTGACCGGCTCTCTCTAGTTGAACCTTGGAACAATTC 1125
QY 1450 CAAACTGTCAGAGGCGGGGCTTGCCTCTCAATCTCCGAGGACCTTTCATCTCAAGC 1509
DB 1126 CAAACTGTCAGAGGCGGGGCTTGCCTCTCAATCTCCGAGGACCTTTCATCTCAAGC 1185
QY 1510 TCAGGCGCCATCCCTCTCTGAGCTCTGACCCCAATTAAGTCCCGAATAAATGAGA 1569
DB 1186 TCAGGCGCCATCCCTCTCTGAGCTCTGACCCCAATTAAGTCCCGAATAAATGAGA 1245
QY 1570 AG 1571
DB 1246 AG 1247

RESULT 7

US-10-295-027-515
; Sequence 515, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 515
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-515

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.4e-132;

[illegible]

```

RESULT 8
US-10-173-999-53
; Sequence 53, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-999-53

Query Match      30.7%      Score 482;   DB 17;   Length 1260;
Best Local Similarity 100.0%      Pctd. No. 3,46-112;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

| | | | | |
|----|-------|--------------|---|-------|
| QY | 10930 | CAGGGTGAATTC | CGGGGGGGCCCTGTGGTCTGCAATGGCTCCCTGGCAGAGGACCTGTGTCCTGG | 11439 |
| Db | 766 | CAGGGTGAATTC | CGGGGGGGCCCTGTGGTCTGCAATGGCTCCCTGGCAGAGGACCTGTGTCCTGG | 825 |
| QY | 11550 | GGAAGATTAC | CCCTTGGGCCCGGACCACAAGACCGGGTGTCTACAGAACTCTTGCAATTC | 12099 |
| Db | 826 | GGAAGATTAC | CCCTTGGGCCCGGACCACAAGACCGGGTGTCTACAGAACTCTTGCAATTC | 885 |
| QY | 1210 | ACCAAGTGAAT | CCAGAAACCATCCAGGCCAACTCTGAGTCATCTCCAGAACTCAGACA | 12656 |
| Db | 886 | ACCAAGTGAAT | CCAGAAACCATCCAGGCCAACTCTGAGTCATCTCCAGAACTCAGACA | 945 |
| QY | 1270 | CCGGATCCCCA | CTGCTGCGAGGACAGCCCTGACATCTCTTTAGAACCTTCATTCCTTC | 13296 |
| Db | 946 | CCGGATCCCCA | CTGCTGCGAGGACAGCCCTGACATCTCTTTAGAACCTTCATTCCTTC | 10059 |
| QY | 1330 | CCAGAGATGTT | GAGAAATGTCATCTCTCAGACCCCTGACCCCATGTCTCTGGACTCAGG | 1389 |
| Db | 1006 | CCAGAGATGTT | GAGAAATGTCATCTCTCTCAGACCCCTGACCCCATGTCTCTGGACTCAGG | 10655 |
| QY | 1390 | GTCGTCTCCCC | CAATTGGGCTGACCGTGTCTCTTAGTGAACCTCTGGAAACAATTC | 1449 |
| Db | 1066 | GTCGTCTCCCC | CAATTGGGCTGACCGTGTCTCTTAGTGAACCTCTGGAAACAATTC | 11225 |
| QY | 1450 | CAAAACTGTCA | GGCGGGGTTGCGTCAATCTCCCTGGGACAATTTCATCTCAAGC | 1509 |
| Db | 1136 | CAAAACTGTCA | GGCGGGGTTGCGTCAATCTCCCTGGGACAATTTCATCTCAAGC | 11859 |
| QY | 1510 | TCAGGAGCCCAT | CTCTCTGTGACGCTGACACCCAAATTAATCCCAAGAAATAACTGAGA | 15656 |
| Db | 1186 | TCAGGAGCCCAT | CTCTCTGTGACGCTGACACCCAAATTAATCCCAAGAAATAACTGAGA | 1245 |
| QY | 1570 | AG | 1571 | |
| Db | 1246 | AG | 1247 | |

```

RESULT 9
US-10-342-887-1411
; Sequence 1411, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Limeley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roderts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1411
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3,4e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1090 CAGGCTGATTTGGGGGCGCTGTGCTGCACATGGCTTCGACGGGACTGCTGCTCTGG 1149

```

Db 766 CAGGGATTCCTGGGGGGGCTGTGTCTGCATGAGTCTCCCTGACGGGAACTGTGTCTCTG 825
QY 1150 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 1209
Db 826 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 885
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAACTCTGTGATCATCCAGAGACTCAGACACA 1269
Db 886 ACCAAGTGAATCCAGGAAACCATCCAGGCCAACTCTGTGATCATCCAGAGACTCAGACACA 945
QY 1270 CCGGATCCCACTGTCTGAGGAGACAGCCCTGACACTCTCTTCAAGACCTCTATCTCTTC 1329
Db 946 CCGGATCCCACTGTCTGAGGAGACAGCCCTGACACTCTCTTCAAGACCTCTATCTCTTC 1005
QY 1330 CCAGAGATGTTGAAGATGTTCTCTCTGAGCCCTGACCCCACTCTCTCTGAGACTCAGG 1389
Db 1006 CCAGAGATGTTGAAGATGTTCTCTCTGAGCCCTGACCCCACTCTCTCTGAGACTCAGG 1065
QY 1390 GTCTGCTTCCCACTGTGAGCTGACCGGTCTCTCTGATTTGAACCTTGGAACCAATTC 1449
Db 1066 GTCTGCTTCCCACTGTGAGCTGACCGGTCTCTCTGATTTGAACCTTGGAACCAATTC 1125
QY 1450 CAAACTGTCCAGGGGGGGGTGTGCTCTCAATCTCTCTGAGGCACTTTCTATCTCAAGC 1509
Db 1126 CAAACTGTCCAGGGGGGGGTGTGCTCTCAATCTCTCTGAGGCACTTTCTATCTCAAGC 1185
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATTAACAGAGA 1569
Db 1186 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATTAACAGAGA 1245
QY 1570 AG 1571
Db 1246 AG 1247

RESULT 10

US-09-739-907-52
Sequence 52, Application US/09739907
Patent No. US20010012889A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/739,907
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
PRIOR FILING DATE: 1998-01-07
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1381
TYPE: DNA
ORGANISM: Homo sapiens
US-09-739-907-52

Query Match 30.7%; Score 482; DB 9; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGATTCCTGGGGGGGCTGTGTCTGCATGAGTCTCCCTGACGGGAACTGTGTCTCTG 1149
Db 852 CAGGGATTCCTGGGGGGGCTGTGTCTGCATGAGTCTCCCTGACGGGAACTGTGTCTCTG 911
QY 1150 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 1209

Db 912 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 971
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAACTCTGTGATCATCCAGAGACTCAGACACA 1269
Db 972 ACCAAGTGAATCCAGGAAACCATCCAGGCCAACTCTGTGATCATCCAGAGACTCAGACACA 1031
QY 1270 CCGGATCCCACTGTCTGAGGAGACAGCCCTGACACTCTCTTCAAGACCTCTATCTCTTC 1329
Db 1032 CCGGATCCCACTGTCTGAGGAGACAGCCCTGACACTCTCTTCAAGACCTCTATCTCTTC 1091
QY 1330 CCAGAGATGTTGAAGATGTTCTCTCTGAGCCCTGACCCCACTCTCTCTGAGACTCAGG 1389
Db 1092 CCAGAGATGTTGAAGATGTTCTCTCTGAGCCCTGACCCCACTCTCTCTGAGACTCAGG 1151
QY 1390 GTCTGCTTCCCACTGTGAGCTGACCGGTCTCTCTGATTTGAACCTTGGAACCAATTC 1449
Db 1152 GTCTGCTTCCCACTGTGAGCTGACCGGTCTCTCTGATTTGAACCTTGGAACCAATTC 1211
QY 1450 CAAACTGTCCAGGGGGGGGTGTGCTCTCAATCTCTCTGAGGCACTTTCTATCTCAAGC 1509
Db 1212 CAAACTGTCCAGGGGGGGGTGTGCTCTCAATCTCTCTGAGGCACTTTCTATCTCAAGC 1271
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATTAACAGAGA 1569
Db 1272 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATTAACAGAGA 1331
QY 1570 AG 1571
Db 1332 AG 1333

RESULT 11

US-09-938-671-52
Sequence 52, Application US/09938671
Publication No. US20040002066A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/938,671
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
PRIOR FILING DATE: 1998-01-07
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1381
TYPE: DNA
ORGANISM: Homo sapiens
US-09-938-671-52

Query Match 30.7%; Score 482; DB 11; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGATTCCTGGGGGGGCTGTGTCTGCATGAGTCTCCCTGACGGGAACTGTGTCTCTG 1149
Db 852 CAGGGATTCCTGGGGGGGCTGTGTCTGCATGAGTCTCCCTGACGGGAACTGTGTCTCTG 911
QY 1150 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 1209
Db 912 GGAGATTACCTCTGTGTGCTCCGGCCCAACAGACGGAGTGTCTACAGAACTCTGCAAGTTC 971
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAACTCTGTGATCATCCAGAGACTCAGACACA 1269

| | | | |
|----|------|--|-------|
| Db | 972 | ACCAAGTGGATCCAGAGAAACCATCCAGGGCAACTCTGATGTCATCCAGAGACTAGACA | 10311 |
| QY | 1270 | CCGGCATCCCCACCTGCTGCAGGGGACAGCCCTGACACTCTTTCAGACCTCTCATCTTC | 13292 |
| Db | 1032 | CCGGCATCCCCACCTGCTGCAGGGGACAGCCCTGACACTCTTTCAGACCTCTCATCTTC | 10910 |
| QY | 1330 | CCAGGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCCTCCGAGCTCAGG | 13889 |
| Db | 1092 | CCAGGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCCTCCGAGCTCAGG | 11510 |
| QY | 1390 | GTCCTACTCCCCACCATTTGGGCTGACCGGTCTCTAGTTGAACCTCTGGGAACAATTC | 14499 |
| Db | 1152 | GTCCTACTCCCCACCATTTGGGCTGACCGGTCTCTAGTTGAACCTCTGGGAACAATTC | 12111 |
| QY | 1450 | CAAAACTGTCCAGGGCGGGGGGTTCGCTCAATCTCCCTGGGGCACTTCACTCTCAAGC | 15099 |
| Db | 1212 | CAAAACTGTCCAGGGCGGGGGGTTCGCTCAATCTCCCTGGGGCACTTCACTCTCAAGC | 12710 |
| QY | 1510 | TCAGGGCCCATCTCTTCTCTGCAAGCTTGACCCCAATTTAGTCCAGAAATPAACTGAGA | 15689 |
| Db | 1272 | TCAGGGCCCATCTCTTCTCTGCAAGCTTGACCCCAATTTAGTCCAGAAATPAACTGAGA | 13310 |
| QY | 1570 | AG | |
| Db | 1332 | AG 1333 | |

```

RESULT 12
US-09-739-907-53
; Sequence 53, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-739-907-53

```

| | Query Match | 30.7% | Score 482 | DB 9 | Length 1439 |
|----|-----------------------|---|--------------------|----------|-------------|
| | Best Local Similarity | 100.0% | Pred. No. 3,6e-132 | | |
| | Matches 482 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1090 | CAGGTTGATTTCTGGGGGGGCTGTGTCGTGCAATGCTCCCTGACAGGACTCGTGTCTCG | 1149 | | |
| Db | 910 | CAGGTTGATTTCTGGGGGGGCTGTGTCGTGCAATGCTCCCTGACAGGACTCGTGTCTCG | 969 | | |
| QY | 1150 | GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGMACTCTGCMAATTC | 1209 | | |
| Db | 970 | GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGMACTCTGCMAATTC | 1029 | | |
| QY | 1210 | ACCAAGTGGATTCACAGAAACCATCCAGGCCCAATCTCTAGTCATCCGAGACTAGACA | 1265 | | |
| Db | 1030 | ACCAAGTGGATTCACAGAAACCATCCAGGCCCAATCTCTAGTCATCCGAGACTAGACA | 1089 | | |
| QY | 1270 | CCGGCATCCCACTGCTGTCAGAGGACAGCCCTGACATCTTTTCAGACCTTCATTCCTTC | 1329 | | |

| | | | |
|----|------|---|------|
| Db | 1090 | CGGGATCCCACTGGTCGACAGGGACAGCCCTGACACTCTCTTTCAGACCCCTCATCTCTTC | 1149 |
| QY | 1330 | CCAGAGATGTTGAGAAATGTTTCATCTCTTCACAGCCCTTGACCCCATGTCCTCTGACTCAGG | 1389 |
| Db | 1150 | CCAGAAGATGTTGAATAATGTTTCATCTTCACAGCCCTTACCCCATAGTCTCTCGACTCAGG | 1209 |
| QY | 1390 | GTCGTCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCTTGGAACAAATTTTC | 1449 |
| Db | 1210 | GTCGTCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCTTGGAACAAATTTTC | 1269 |
| QY | 1450 | CAAAACGTGTCAGAGGGCGGGGGTTCGGTCTCAATTCCTCGTGGGACATTTCATCCTCAAGC | 1509 |
| Db | 1270 | CAAAACGTGTCAGAGGGCGGGGGTTCGGTCTCAATTCCTCGTGGGACATTTCATCCTCAAGC | 1329 |
| QY | 1510 | TCAGGGCCCATCCCTTCTCTGACAGCTGACCCCAAAATTTAGTCCACAGAAATTAACCTGAGA | 1569 |
| Db | 1330 | TCAGGGCCCATCCCTTCTCTGACAGCTGACCCCAAAATTTAGTCCACAGAAATTAACCTGAGA | 1389 |
| QY | 1570 | AG AG 1571 | |
| Db | 1390 | AG 1391 | |

```

US-09-938-671-53      RESULT 13
; Sequence 53, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/938,671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,652
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-53

```

| Query Match | 30.7% | Score 482 | DB 11 | Length 1439 |
|-----------------------|----------------|--|----------|-------------|
| Best Local Similarity | 100.0% | Pred. No. 3.6e-132 | | |
| Matches 482 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1090 | CAGGGTATTCCTGGGGGGCCCTGTGGTCTGCAGTGGCTCCCTGCAGGGACTGATCTCTGG | 1149 | |
| Db | 910 | CAGGGTATTCCTGGGGGGCCCTGTGGTCTGCAGTGGCTCCCTGCAGGGACTGATCTCTGG | 969 | |
| QY | 1150 | GGAGATTACCTTGTGCCCCGGGCCACAGACCGGGTGTCTACACGAACCTTGTCAATTTC | 1209 | |
| Db | 970 | GGAGATTACCTTGTGCCCCGGGCCACAGACCGGGTGTCTACACGAACCTTGTCAATTTC | 1029 | |
| QY | 1210 | ACCAAGGAGATCCAGGAAACCATCCAGGCCAATCCTGAGATCAATCCAGAGACTCAGACA | 1269 | |
| Db | 1030 | ACCAAGTGGATCCAGAAACCATCCAGGCCAATCCTGAGATCAATCCAGAGACTCAGACA | 1089 | |
| QY | 1270 | CCGGCATTCGCCACTGCTGTCAGGGACAGCCCTTGACATCTCTTTTCAGACCCCTCATTTCTTC | 1329 | |
| Db | 1090 | CCGGCATTCGCCACTGCTGTCAGGGACAGCCCTTGACATCTCTTTTCAGACCCCTCATTTCTTC | 1149 | |
| QY | 1330 | CCAGAGATGTTAGAGATGTTCAATCTTCCAGCCCTGAGCCCAATGTTCTCTGGACTCAGG | 1389 | |

Db 1150 CCAGAGATGTGAGATGTTCACTCTCCAGCCCTGAGCCCATGTCCTGAGACTCAGG 1209
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1449
Db 1210 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1269
QY 1450 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1509
Db 1270 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1329
QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1569
Db 1330 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1389
QY 1570 AG 1571
Db 1390 AG 1391

RESULT 14

US-10-723-860-5032
; Sequence 5032, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Azile, Natasha
; APPLICANT: Gishburg, Wendy M.
; APPLICANT: Zlonchik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPDS01
; CURRENT FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5032
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5032

Query Match 30.7%; Score 482; DB 18; Length 1481;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGAGACTGCTGTCTGG 1149
Db 894 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGAGACTGCTGTCTGG 953
QY 1150 GGAAGATTAACCTTGTGCCCCGGGCCAAAGACCGGGGTGTACACGAACCTTGCAAGTTC 1209
Db 954 GGAAGATTAACCTTGTGCCCCGGGCCAAAGACCGGGGTGTACACGAACCTTGCAAGTTC 1013
QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1269
Db 1014 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1073
QY 1270 CCGGATATCCCACTGCTGCAAGGAGACGCTTGACACTCTCTTCAAGACCTTCAATCTTTC 1329
Db 1074 CCGGATATCCCACTGCTGCAAGGAGACGCTTGACACTCTCTTCAAGACCTTCAATCTTTC 1133
QY 1330 CCAGAGATGTGAGAAATGTTCACTCTCAGGCCCTGAGCCCAATGTCCTGAGACTCAGG 1389
Db 1134 CCAGAGATGTGAGAAATGTTCACTCTCAGGCCCTGAGCCCAATGTCCTGAGACTCAGG 1193
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1449
Db 1194 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1253
QY 1450 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1509
Db 1254 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1313

QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1569
Db 1314 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1373
QY 1570 AG 1571
Db 1374 AG 1375

RESULT 15

US-09-739-907-37
; Sequence 37, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022Pl
; CURRENT FILING DATE: 2000-12-20
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,c,g, or c
US-09-739-907-37

Query Match 30.7%; Score 482; DB 9; Length 1516;
Best Local Similarity 100.0%; Pred. No. 3.7e-112;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGAGACTGCTGTCTGG 1149
Db 977 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGAGACTGCTGTCTGG 1036
QY 1150 GGAAGATTAACCTTGTGCCCCGGGCCAAAGACCGGGGTGTACACGAACCTTGCAAGTTC 1209
Db 1037 GGAAGATTAACCTTGTGCCCCGGGCCAAAGACCGGGGTGTACACGAACCTTGCAAGTTC 1096
QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1269
Db 1097 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1156
QY 1270 CCGGATATCCCACTGCTGCAAGGAGACGCTTGACACTCTCTTCAAGACCTTCAATCTTTC 1329
Db 1157 CCGGATATCCCACTGCTGCAAGGAGACGCTTGACACTCTCTTCAAGACCTTCAATCTTTC 1216
QY 1330 CCAGAGATGTGAGAAATGTTCACTCTCAGGCCCTGAGCCCAATGTCCTGAGACTCAGG 1389
Db 1217 CCAGAGATGTGAGAAATGTTCACTCTCAGGCCCTGAGCCCAATGTCCTGAGACTCAGG 1276
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1449
Db 1277 GTCGTCTTCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCTGGAAACAATTTC 1336
QY 1450 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1509
Db 1337 CAAAACGTGTCAGGGGGGGGGTTCAGTCTCAATCTTCCCTGGGCACTTTCATCTCAAGC 1396

| | | | |
|----|------|---|------|
| QY | 1510 | TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTAGTCCCGAGAAATAAAGTGAGA | 1569 |
| Db | 1397 | TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTAGTCCCGAGAAATAAAGTGAGA | 1456 |
| QY | 1570 | AG 1571 | |
| Db | 1457 | AG 1458 | |

Search completed: February 27, 2005, 04:52:32
Job time : 850.613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 20:26:59 ; Search time 4923.95 Seconds
(without alignments)
12144.524 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 agggagagcttcggggaactga.....cccaagaataactgagaag 1571

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| C 1 | 482 | 30.7 | 539 | 2 | BF513278 UT-H-BM1- |
| C 2 | 482 | 30.7 | 586 | 7 | W73140 zds5611.B1 |
| C 3 | 475.8 | 30.3 | 809 | 4 | BG680075 BG680075 |
| C 4 | 455 | 29.0 | 478 | 1 | AA682032 AA682032 |
| C 5 | 443.8 | 28.2 | 802 | 4 | BG682309 BG682309 |
| C 6 | 440 | 28.0 | 953 | 2 | BE745465 BE745465 |
| C 7 | 431.8 | 27.5 | 666 | 4 | BM842155 BM842155 |
| C 8 | 417 | 26.5 | 645 | 1 | AI002163 AI002163 |
| C 9 | 383.6 | 24.4 | 677 | 2 | BE898385 BE898385 |
| C 10 | 381.8 | 24.3 | 453 | 3 | NB0762 NB0762 |
| C 11 | 355.8 | 22.6 | 363 | 2 | AM105502 AM105502 |
| C 12 | 348 | 22.2 | 590 | 2 | BM840511 BM840511 |
| C 13 | 333.4 | 21.2 | 577 | 4 | BM838406 BM838406 |
| C 14 | 322.8 | 20.5 | 450 | 7 | W68361 W68361 |
| C 15 | 317 | 20.2 | 388 | 7 | W68496 W68496 |
| C 16 | 314 | 20.0 | 334 | 2 | BF514439 BF514439 |
| C 17 | 291 | 18.5 | 533 | 4 | BM840509 BM840509 |
| C 18 | 278.8 | 17.7 | 523 | 4 | BM841293 BM841293 |
| C 19 | 277.2 | 17.6 | 320 | 2 | AM380655 AM380655 |
| C 20 | 276.6 | 17.6 | 411 | 2 | AM801647 AM801647 |
| C 21 | 273 | 17.4 | 512 | 4 | BM830263 BM830263 |
| C 22 | 270.4 | 17.2 | 526 | 4 | BM841697 BM841697 |
| C 23 | 245 | 15.6 | 498 | 7 | W73168 W73168 |
| C 24 | 229.8 | 14.6 | 420 | 6 | CB298043 CB298043 |

| | | | | | |
|------|-------|------|------|---|----------|
| 25 | 229.2 | 14.6 | 422 | 4 | BG398353 |
| 26 | 227.8 | 14.5 | 947 | 4 | BG680848 |
| C 27 | 227.6 | 14.5 | 479 | 2 | AM576251 |
| C 28 | 227.6 | 14.5 | 486 | 1 | AI755214 |
| C 29 | 227.6 | 14.5 | 726 | 5 | BQ005946 |
| C 30 | 227.6 | 14.5 | 726 | 7 | CN395303 |
| C 31 | 227.6 | 14.5 | 961 | 3 | BU197099 |
| C 32 | 227.6 | 14.5 | 4544 | 5 | HSMB0369 |
| C 33 | 226 | 14.4 | 359 | 2 | AM500684 |
| C 34 | 226 | 14.4 | 476 | 1 | AI754567 |
| C 35 | 226 | 14.4 | 481 | 1 | AI754105 |
| C 36 | 226 | 14.4 | 740 | 7 | CN395308 |
| C 37 | 226 | 14.4 | 977 | 6 | CD516130 |
| C 38 | 226 | 14.4 | 1092 | 4 | BM480074 |
| C 39 | 226 | 14.4 | 1547 | 3 | CR621085 |
| C 40 | 226 | 14.4 | 1672 | 3 | CR615950 |
| C 41 | 223.4 | 14.2 | 3341 | 3 | CR749847 |
| C 42 | 223 | 14.2 | 744 | 4 | BG743962 |
| C 43 | 222.4 | 14.2 | 715 | 2 | BF574982 |
| C 44 | 222 | 14.1 | 352 | 1 | AI249688 |
| C 45 | 221.2 | 14.1 | 649 | 4 | BG399464 |

ALIGNMENTS

RESULT 1
BF513278/c
LOCUS
DEFINITION
UI-H-BM1-amo-e-03-0-UI.B1 NCI CGAP_Sub7 Homo sapiens cDNA clone
IMAGE:3070564 3', mRNA sequence.
ACCESSION
BF513278
VERSION
BF513278.1 GI:11598457
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 539)
NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
NCI CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Seq primer: M13 Forward
POLYVA=Yes.

FEATURES

Source

Location/Qualifiers
1..539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3070564"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Sub7"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7
is a subcloned library derived from NCI CGAP Sub6. The
NCI CGAP Sub7 library had 12 million recombinants. A
single-stranded DNA preparation of NCI CGAP Sub6 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI CGAP_Kids pool 1 LLAM
3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)
1322376-1323911, 1456008-1456775, 1500552-1502855);
NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE Clones) 1323912-1325831, 1471368-1472803.

1492104-1492325) ; NCI CGAP L5 pool 1 L1AM 3575-3582,
3851-3854 (IMAGE Clonoids L414990-1417991,
1520904-1522439) ; NCI CGAP C64 pool 1 L1AM 3164-3167,
3716-3720, 3733-3735 (IMAGE Clonoids
1257096-1258631, 1465064-1470983, 1475592-1476743) ;
NCI CGAP Pr22 pool 1 L1AM 2457-2459, 2758-2759, 3062-3068
(IMAGE Clonoids 985608-986759, 1101192-1101959,
1217928-1220615) ; NCI CGAP C010 pool 1 L1AM 2644-2653,
2871-2872 (IMAGE Clonoids
1057416-1061255, 1144586-1145531). (6% of the driver
population), plus a pool of 3,840 arrayed clones from
NCI CGAP Sub1 (IMAGE Clonoids 2708616-2710535) and
NCI CGAP Sub2 (IMAGE Clonoids 2710536-2712455) (4% of
the driver population), plus a pool of 11,136 clones from
NCI CGAP Sub3 (IMAGE Clonoids 2712456-2723591) (10% of
the driver population), plus a pool of 5,472 clones from
NCI CGAP Sub4 (IMAGE Clonoids 2723592-2723326) (40% of the
driver population), plus a pool of 4032 clones from
NCI CGAP Sub6 (IMAGE Clonoids 2728969-2733190) (40% of the
driver population). Subtraction was performed as
previously described [Bonaldi, Lennon & Soares (1996) :
Normalization and Subtraction: Two Approaches To
Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_T15SUB-1ung
TAG_L1B=NCI CGAP-Lu13
TAG_SEQ=GCCG"

ORIGIN

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 30.7%; | Score 482; | DB 2; | Length 539; |
| Best Local Similarity | 100.0%; | Pred. No. 4.8e-65; | | |
| Matches 482; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|------|---|------|
| QY | 1090 | TAGGGTATATTCGGGGGGGGCCGTGTCTGCAGTAATGCTCCCTGCAGGGACTGTGTCCGG | 1149 |
| Db | 508 | CAGGTGTATTCGGGGGGCCGTGTGTCTGCAGTAATGCTCCCTGCAGGGACTGTGTCCGG | 449 |
| QY | 1150 | GGAATTAACCTTGTGTCCGGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAGTTTC | 1209 |
| Db | 448 | GGAGATTACCTTGTGTCCGGGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAGTTTC | 389 |
| QY | 1210 | ACCAATGTGATTCGAGGAACCATTCAGGCCCAATCTCTAGTCATCTCCAGACTCAGACA | 1265 |
| Db | 388 | ACCAATGTGATTCGAGGAACCATTCAGGCCCAATCTCTAGTCATCTCCAGACTCAGACA | 329 |
| QY | 1270 | CCGGCATCCCAACCTGTCTCAGGGACAGCCCTGACACTCTCTTTCAGACCTTCATTCCTTC | 1329 |
| Db | 328 | CCGGCATCCCAACCTGTCTCAGGGACAGCCCTGACACTCTCTTTCAGACCTTCATTCCTTC | 269 |
| QY | 1330 | CCAAGATGTGAAATGTTCATCTCTCAGACCCCTGACCCCATGTCTCTCTGACTCAGG | 1389 |
| Db | 268 | CCAAGATGTGAAATGTTCATCTCTCAGACCCCTGACCCCATGTCTCTCTGACTCAGG | 209 |
| QY | 1390 | GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCTGGAAACAATTC | 1449 |
| Db | 208 | GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCTGGAAACAATTC | 149 |
| QY | 1450 | CAAAATGTCCAGGGCCGGGGGTTGCTGTCAATCTCCTGGGGCACTTTACCTCAAGC | 1509 |
| Db | 148 | CAAAATGTCCAGGGCCGGGGGTTGCTGTCAATCTCCTGGGGCACTTTACCTCAAGC | 89 |
| QY | 1510 | TCAGGGCCCATCCCTTCTCTGACGTGACCCAAATTTAGTCCAGAAATTAACCTGAGA | 1569 |
| Db | 88 | TCAGGGCCCATCCCTTCTCTGACGTGACCCAAATTTAGTCCAGAAATTAACCTGAGA | 29 |
| QY | 1570 | AG 1571 | |
| Db | 28 | AG 27 | |

| | | | | |
|------------|--|--------|------|--------|
| RESULT 2 | | | | |
| W73140/c | | | | |
| LOCUS | W73140 | 586 bp | mRNA | linear |
| DEFINITION | zdfs5e11.g1 Soares_fetal_heart_NbHn19w Homo sapiens cDNA clone | | | |

IMAGE:344588.3', similar to PIR:A53968 A53968 serine proteinase SCCE
 precursor - human ;, mRNA sequence.

ACCESSION W73140
 VERSION W73140.1 GI:1383275
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 586)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lemon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P.,
 Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.

The WashU-Merck EST Project
 Unpublished (1995)

COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 761 Std Error: 0.00
 Seq primer: ETRprimer
 High quality sequence stop: 428.
 Location/Qualifiers

FEATURES
 source 1..586

ORIGIN

| | | | | |
|-----------------------|---------|--------------------|-------|-------------|
| Query Match | 30.7%; | Score 482; | DB 7; | Length 586; |
| Best Local Similarity | 100.0%; | Pred. No. 4.6e-65; | | |

| | | | |
|----|------|---|-------|
| Qy | 1090 | CAGGATTAATTCGTGGGGGGGCTCTGTGTCTTGCAATAGGCTCCCTGACAGGGACATCTGTCCCTGG | 11459 |
| Db | 501 | CAGGATTAATTCGTGGGGGGGCTCTGTGTCTTGCAATAGGCTCCCTGACAGGGACATCTGTCTCTGG | 442 |
| Qy | 1150 | GGAGATTACCTTGTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACTCTGCAAGTTC | 1209 |
| Db | 441 | GGAGATTACCTTGTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACTCTGCAAGTTC | 382 |
| Qy | 1210 | ACCAAGTGATTCACAGAAACCATTCAGGCGCAACTCTTGAGTCATCCAGGACTCAGACACA | 1265 |
| Db | 381 | ACCAAGTGATTCACAGAAACCATTCAGGCGCAACTCTTGAGTCATCCAGGACTCAGACACA | 322 |
| Qy | 1270 | CCGGCATCCCCACTGTGTGACAGGACAGCCCTGACATCTCTTTCAAGCCCTCATTTCTCTTC | 1329 |
| Db | 321 | CCGGCATCCCCACTGTGTGACAGGACAGCCCTGACATCTCTTTCAAGCCCTCATTTCTCTTC | 262 |

QY 1330 CCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTGAGCTCAG 1389
DB 261 CCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTGAGCTCAG 202
QY 1390 GTCTGCTTCCCGACATTTGGGCTGACCGCTCTCTCTAGTTGAACCTTGGAAATTTTC 1449
DB 201 GTCTGCTTCCCGACATTTGGGCTGACCGCTCTCTCTAGTTGAACCTTGGAAATTTTC 142
QY 1450 CAAATCTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTCTGAGGACCTTTTCTCTCAGC 1509
DB 141 CAAATCTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTCTGAGGACCTTTTCTCTCAGC 82
QY 1510 TCAGGCGCCATCTCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTAAGTGA 1569
DB 81 TCAGGCGCCATCTCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTAAGTGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 3
BG680075 809 bp mRNA linear EST 01-MAY-2001
LOCUS 602628224F1 NCI_CGAP_Skin4 Homo sapiens cDNA clone IMAGE:4753372 5',
DEFINITION mRNA sequence.
ACCESSION BG680075 GI:13911472
VERSION BG680075.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS NIH (bases 1 to 809)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10612 row: 0 column: 05
High quality sequence stop: 707.
Location/Qualifiers
1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753372"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skin4"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: Nci1;
Site 2: Sal1; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 30.3%; Score 475.8; DB 4; Length 809;
Best Local Similarity 99.6%; Pired. No. 3.8e-64;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1084 CTCCTGAGGATGATTTCTGAGGGGCTGCTGATGATGCTCCCGAGGAGCTCTG 1143
DB 219 CTCCTGAGGATGATTTCTGAGGGGCTGCTGATGATGCTCCCGAGGAGCTCTG 278
QY 1144 TCCCTGGGAGATTACCTTGTGCGCGGCCAAGACCGGGTCTTACAGAACTCTGC 1203

DB 279 TCTGAGGAGATTACCTTGTGCGCGGCCAAGACCGGGTCTTACAGAACTCTGC 338
QY 1204 AAGTTACCAATGATTCAGAGAAACATCCAGGCCAATCTCTGATGATCTCCAGAGATC 1263
DB 339 AAGTTACCAATGATTCAGAGAAACATCCAGGCCAATCTCTGATGATCTCCAGAGATC 398
QY 1264 AGCAGACCGGATCCCACTGCTGAGGAGAGAGCCCTGACACTCTCTTTCAGACCTCAT 1323
DB 399 AGCAGACCGGATCCCACTGCTGAGGAGAGAGCCCTGACACTCTCTTTCAGACCTCAT 458
QY 1324 TCTTCCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTCTGA 1383
DB 459 TCTTCCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTCTGA 518
QY 1384 CTCAGGCTGCTTCCCGACATTTGGGCTGACCGCTCTCTCTAGTTGAACCTTGGAAAC 1443
DB 519 CTCAGGCTGCTTCCCGACATTTGGGCTGACCGCTCTCTCTAGTTGAACCTTGGAAAC 578
QY 1444 AATTTCCAAACTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGACCTTTTCATCC 1503
DB 579 AATTTCCAAACTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGACCTTTTCATCC 638
QY 1504 TCAAGCTCAGGCGCCATCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTA 1562
DB 639 TCAAGCTCAGGCGCCATCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTA 697

RESULT 4
AA862032 478 bp mRNA linear EST 13-APR-1999
LOCUS 0146609.s1 NCI_CGAP_HN3 Homo sapiens cDNA clone IMAGE:1485736 3',
DEFINITION similar to SW_SCE_HUMAN P49862 STRATUM CORNEUM CHYMOTRYPTIC ENZYME
PRECURSOR ;, mRNA sequence.
ACCESSION AA862032 GI:2954511
VERSION AA862032.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NCI/NIDH-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/
AUTHORS NCI/NIDH-CGAP (bases 1 to 478)
JOURNAL National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: John Emley, M.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1647 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 233.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1485736"
/tissue_type="squamous cell carcinoma from base of tongue"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_HN3"
/note="Organ: tongue; Vector: Bluescript SK-; Site: 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.0 kb. 5' adaptor sequence:
5' GAATTCGAGCAG 3' 3' adaptor sequence: 5'
(GA)10ACTAGTCTCAGATTTTTTTTTTTTTTTT 3'."

ORIGIN

Query Match 29.0%; Score 455; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 7.5e-61;
Matches 466; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1105 GGGCCCTGTGTCTGCAATGGCTTCCCTGCAAGGACTGTGTCTCTGGGAGATTACCTTGT 1164
DB 478 GGGCCCTGTGTCTGCAATGGCTTCCCTGCAAGGACTGTGTCTCTGGGAGATTACCTTGT 420
QY 1165 GCCCGGCCCAAGACCGGGGTGTCTACAGAACTCTGCAAGTTCACCAAGTGGATCCAG 1224
DB 419 GCCCGGCCCAAGACCGGGGTGTCTACAGAACTCTGCAAGTTCACCAAGTGGATCCAG 360
QY 1225 GAACCATTCAGGCACTCTGAGTCATCCAGACTGACACACCGGATCCCACT 1284
DB 359 GAACCATTCAGGCACTCTGAGTCATCCAGACTGACACACCGGATCCCACT 300
QY 1285 GCTGCAAGGACAGCCCTGACACTCTCTTTCAGACCTTCATTCCTCCAGATGTTGAGA 1344
DB 299 GCTGCAAGGACAGCCCTGACACTCTCTTTCAGACCTTCATTCCTCCAGATGTTGAGA 240
QY 1345 ATGTTCACTCTCCAGACCCCTGACCCCATGTCTCTGACTGAGGCTGTCTCCCCAC 1404
DB 239 ATGTTCACTCTCCAGACCCCTGACCCCATGTCTCTGACTGAGGCTGTCTCCCCAC 180
QY 1405 ATGAGGCTGACCGTGTCTCTGATTGAACCTGAGGACAAATTCMAAATGTCCAGG 1464
DB 179 ATGAGGCTGACCGTGTCTCTGATTGAACCTGAGGACAAATTCMAAATGTCCAGG 120
QY 1465 CCGGGGTTCCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGCTCAGGCCCCATCCT 1524
DB 119 CCGGGGTTCCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGCTCAGGCCCCATCCT 60
QY 1525 TCTCTGACCTGACCCCAATTTAGTCCAGAAATTAACGTAGAG 1571
DB 59 TCTCTGACCTGACCCCAATTTAGTCCAGAAATTAACGTAGAG 13

RESULT 5
BG682309 802 bp mRNA 1linear EST 01-MAY-2001
LOCUS 60269926F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',
DEFINITION mRNA Sequence.
ACCESSION BG682309
VERSION BG682309.1 GI:13913706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10615 row: d column: 08
High quality sequence stop: 802.
Location/Qualifiers
1..802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754263"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NCI_CGAP_Skn4"

FEATURES
source

/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP library."

Query Match 28.2%; Score 443.8; DB 4; Length 802;
Best Local Similarity 99.3%; Pred. No. 3.5e-59;
Matches 456; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1090 CAGGATGATTTCTGGGAGGCTGTGTCTGCAATGCTCTCCAGAGACTGTCTCTGG 1149
DB 333 CAGGATGATTTCTGGGAGGCTGTGTCTGCAATGCTCTCCAGAGACTGTCTCTGG 392
QY 1150 GGAAGATTACCTTGCCCGGCCCAAGACCGGGGTCTACAGAACTCTGCAAGTTC 1209
DB 393 GGAAGATTACCTTGCCCGGCCCAAGACCGGGGTCTACAGAACTCTGCAAGTTC 452
QY 1210 ACCAAGTGATCCAGAAACCATCCAGGCCAATCTCTGATCATCCAGACTCAGACACA 1269
DB 453 ACCAAGTGATCCAGAAACCATCCAGGCCAATCTCTGATCATCCAGACTCAGACACA 512
QY 1270 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTTCAGACCTCATTCCTTC 1329
DB 513 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTTCAGACCTCATTCCTTC 572
QY 1330 CCAGAGATTTGAGAAATGTAATCTCTCAGCCCTGACCCCATGCTCTGAGACTCAGG 1389
DB 573 CCAGAGATTTGAGAAATGTAATCTCTCAGCCCTGACCCCATGCTCTGAGACTCAGG 632
QY 1390 GTCTCTTCCCCACA-TTGGGCTGACCGTGTCTCTGATTGAACCTGAGAACATTT 1448
DB 633 GTCTCTTCCCCACAATTTGGGCTGACCGTGTCTCTGATTGAACCTGAGAACATTT 692
QY 1449 CCAAAACTGTCCAGGGCCGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCATCTCAAG 1508
DB 693 CCAAAACTGTCCAGGGCCGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCATCTCAAG 752
QY 1509 CTCAGGGCCCAATCCCTTCTCTGACGCTGACCCCAATT 1547
DB 753 CTCAGGGCCCAATCCCTTCTCTGACGCTGACCCCAATT 791

RESULT 6
BE745465 953 bp mRNA 1linear EST 15-SEP-2000
LOCUS 601579834F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928640 5',
DEFINITION mRNA Sequence.
ACCESSION BE745465
VERSION BE745465.1 GI:10159457
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM760 row: c column: 09
High quality sequence stop: 685.
Location/Qualifiers
1..953
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

ORIGIN

Query Match 28.0%; Score 440; DB 2; Length 953;
Best Local Similarity 99.6%; Pred. No. 1.3e-56;
Matches 462; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

/db_xref="taxon:9606"
/clone="IMAGE:392864.0"
/issue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NH_MGC_9"
/note="Organ: ovary; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

1090 CAGGATGATCTGGGGGCGCTGTGCAATGCTCCCTGACAGGACTCGTCTCG 1149
|||
229 CAGGATGATCTGGGGGCGCTGTGCAATGCTCCCTGACAGGACTCGTCTCG 288
|||
1150 GGAATTAACCTTGTGCGCCGACACAGACCGGGTGTACACGAACCTCTGCAATTG 1209
|||
289 GGAATTAACCTTGTGCGCCGACACAGACCGGGTGTACACGAACCTCTGCAATTG 348
|||
1210 ACCAAGTGATCCAGAAACATTCAGGCCCACTCTGAGTATCCAGAGCTCAGACA 1269
|||
349 ACCAAGTGATCCAGAAACATTCAGGCCCACTCTGAGTATCCAGAGCTCAGACA 408
|||
1270 CCGGCAATCCGACCTGTGAGGAGGACGCTGACACTCCTTTCAGACCTCAATCTTC 1329
|||
409 CCGGCAATCCGACCTGTGAGGAGGACGCTGACACTCCTTTCAGACCTCAATCTTC 468
|||
1330 CCAGAGATGTTGAAATGTTCAATCTCTCAAGCCCTGACACCCATGCTCTCTGAG 1389
|||
469 CCAGAGATGTTGAAATGTTCAATCTCTCAAGCCCTGACACCCATGCTCTCTGAG 528
|||
1390 GTCTGCTTCCCTCCACATGAGGCTGACCGTGTCTCTAGTTGAAACCTCTGGAACATTTTC 1449
|||
529 GTCTGCTTCCCTCCACATGAGGCTGACCGTGTCTCTAGTTGAAACCTCTGGAACATTTTC 587
|||
1450 CAAACTGTCCAGGGGGGGGGTGTGCTCAATCTCCCTGGGACCTTATCTCTCAGC 1509
|||
588 CAAACTGTCCAGGGGGGGGGTGTGCTCAATCTCCCTGGGACCTTATCTCTCAGC 647
|||
1510 TCAGGGCCCATCCCTCTCTGACAGCTGACCCCAATTTAGTCC 1553
|||
648 TCAGGGCCCATCCCTCTCTGACAG-TCCTGACCCCAATTTAGTCC 690
|||

RESULT 7
BM842155 666 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0119498 S12SNU216 Homo sapiens cDNA clone S12SNU216-56-H11 5',
DEFINITION mRNA sequence.
ACCESSION BM842155
VERSION BM842155.1 GI:19198564
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 666)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Institute of Bioscience & Biotechnology

52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krdb.re.kr
Plate: 56 row: H column: 11
High quality sequence stop: 666.
Location/Qualifiers
1. 666

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-56-H11"
/sex="F"
/issue_type="lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pcns; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 27.5%; Score 431.8; DB 4; Length 666;
Best Local Similarity 99.5%; Pred. No. 2.7e-57;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1090 CAGGATGATCTGGGGGCGCTGTGCAATGCTCCCTGACAGGACTCGTCTCG 1149
|||
232 CAGGATGATCTGGGGGCGCTGTGCAATGCTCCCTGACAGGACTCGTCTCG 291
|||
1150 GGAATTAACCTTGTGCGCCGACACAGACCGGGTGTACACGAACCTCTGCAATTG 1209
|||
292 GGAATTAACCTTGTGCGCCGACACAGACCGGGTGTACACGAACCTCTGCAATTG 351
|||
1210 ACCAAGTGATCCAGAAACATTCAGGCCCACTCTGAGTATCCAGAGCTCAGACA 1269
|||
352 ACCAAGTGATCCAGAAACATTCAGGCCCACTCTGAGTATCCAGAGCTCAGACA 411
|||
1270 CCGGCAATCCGACCTGTGAGGAGGACGCTGACACTCCTTTCAGACCTCAATCTTC 1329
|||
412 CCGGCAATCCGACCTGTGAGGAGGACGCTGACACTCCTTTCAGACCTCAATCTTC 471
|||
1330 CCAGAGATGTTGAAATGTTCAATCTCTCAAGCCCTGACACCCATGCTCTCTGAG 1389
|||
472 CCAGAGATGTTGAAATGTTCAATCTCTCAAGCCCTGACACCCATGCTCTCTGAG 531
|||
1390 GTCTGCTTCCCTCCACATGAGGCTGACCGTGTCTCTAGTTGAAACCTCTGGAACATTTTC 1449
|||
532 GTCTGCTTCCCTCCACATGAGGCTGACCGTGTCTCTAGTTGAAACCTCTGGAACATTTTC 591
|||
1450 CAAACTGTCCAGGGGGGGGGTGTGCTCAATCTCCCTGGGACCTTATCTCTCAGC 1509
|||
592 CAAACTGTCCAGGGGGGGGGTGTGCTCAATCTCCCTGGGACCTTATCTCTCAGC 651
|||
1510 TCAGGGCCCATCCCT 1524
|||
652 TCAGGGCCCATCCCT 666
|||

| LOCUS | AI002163/C | 645 bp | mRNA | linear | EST 27-AUG-1998 |
|---------------------------|--|--------|------|--------|-----------------|
| DEFINITION | ot42h05.s1 Soares_testis_NHT Homo sapiens CDNA IMAGE:1619481 | | | | |
| ACCESSION | AI002163 | | | | |
| KEYWORDS | sequence. | | | | |
| SOURCE | AI002163 | | | | |
| ORGANISM | AI002163.1 GI:3202497 | | | | |
| REFERENCE | EST. | | | | |
| AUTHORS | Homo sapiens (human) | | | | |
| TITLE | Unpublished (1997) | | | | |
| JOURNAL | Contact: Robert Strusberg, Ph.D. | | | | |
| COMMENT | Email: cgsaps-remail.nih.gov | | | | |
| | CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima | | | | |
| | Bonaldo, Ph.D. | | | | |
| | CDNA Library Arrayed by: Greg Lennon, Ph.D. | | | | |
| | DNA Sequencing by: Washington University Genome Sequencing Center | | | | |
| | Clone distribution: NCI-CGAP clone distribution information can be | | | | |
| | found through the I.M.A.G.E. Consortium/LLNL at: | | | | |
| | www.bio.lnl.gov/bbfp/image/image.html | | | | |
| | Insert Length: 1042 Std Error: 0.00 | | | | |
| | Seq primer: -40ml3 fwd. ET from Amerham | | | | |
| | High quality sequence stop: 303. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| Source | 1..645 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="mRNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| | /clone="IMAGE:1619481" | | | | |
| | /sex="male" | | | | |
| | /lab_host="DH10B" | | | | |
| | /clone_lib="Soares_testis_NHT" | | | | |
| | /note="Vector: pUT3D-Pac (Pharmacia) with a modified | | | | |
| | polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA | | | | |
| | was prepared from mRNA obtained from Clontech | | | | |
| | Laboratories, Inc., and primed with a Not I - oligo(dT) | | | | |
| | primer [5', | | | | |
| | TGTTACCAATCTGAAGTGGAGCGCGCCCAATTCTTTTCTTTT 3']. | | | | |
| | Double-stranded cDNA was ligated to Eco RI adaptors | | | | |
| | (Pharmacia), digested with Not I and cloned into the Not I | | | | |
| | and Eco RI sites of the modified pUT3 vector. Library | | | | |
| | went through one round of normalization to Cot5, and was | | | | |
| | constructed by Bento Soares and M. Fatima Bonaldo." | | | | |
| ORIGIN | | | | | |
| Query Match | 26.5%; Score 417; DB 1; Length 645; | | | | |
| Best Local Similarity | 94.9%; Pred. No. 5,3e-55; | | | | |
| Matches 464; Conservative | 0; Mismatches 20; Indels 5; Gaps 3; | | | | |
| 1084 | CTCTGCAAGGATTTCTGGGGGCGCTGTGCTTCAAATGCTCCCTGCAGGACTGTG | 1143 | | | |
| 499 | CTCCGCAAGGATTTCTGGGGGCGCTGTGCTTCAAATGCTCCCTGCAGGACTGTG | 441 | | | |
| 1144 | TCTCGGGGAGATTACCTTTGTCGCCGGGCCAAAGACGGGGTCTTACAGAACTCTGC | 1203 | | | |
| 440 | --CTGGGAGATTACCTTTGT-CCGGGCCAAAGACGGGGTCTTACAGAACTCTGC | 384 | | | |
| 1204 | AAGTTACCAAGTGATCCAGGAACATTCAGGCACTCCTGAGTCATCCAGAGATC | 1263 | | | |
| 383 | AAGTTACCAAGTGATCCAGGAACATTCAGGCACTCCTGAGTCATCCAGAGATC | 324 | | | |
| 1264 | AGCACACCGGCAATCCCACTGCTGCAAGGACAGCCCTGACATCTCTTCAGACCTCAT | 1323 | | | |
| 323 | AGCACACCGGCAATCCCACTGCTGCAAGGACAGCCCTGACATCTCTTCAGACCTCAT | 264 | | | |
| 1324 | TCTTCCAGAGATTTAGATATGTATCTCTCAAGCCCTGACCCCATGTCTCTGGA | 1383 | | | |

| Db | Accession | Source | Organism | Reference | Authors | Title | Journal | Comment |
|-----------------------|--|--|--------------|------------------------------|---------|-------|---------|---------|
| Db | 263 | TCCTTCCAGAGATGTTGAGATGTTTATCTCTCCAGCCCTGACCCCAATGTCCTGGA | 204 | | | | | |
| Qy | 1384 | CTCAGGGCTGCTTCCCTCCACATTTGGGCTGACCGCTGCTCTTATGTTGAACCTTGGGAC | 1443 | | | | | |
| Db | 203 | CTCAGGGCTGCTTCCCTCCACATTTGGGCTGACCGCTGCTCTTATGTTGAACCTTGGGAC | 144 | | | | | |
| Qy | 1444 | AATTTCCAAACTGTCCAGAGGGGGGGTGGTGTCAATCTCCCTGGGGCACTTCAATCC | 1503 | | | | | |
| Db | 143 | AATTTCCAAACTGTCCAGAGGGGGGGTGGTGTCAATCTCCCTGGGGCACTTCAATCC | 84 | | | | | |
| Qy | 1504 | TCAAGCTCAGAGGCCATCTCTTCTCTGACGCTGACCCAAATTAGT-CCAGAAATTA | 1562 | | | | | |
| Db | 83 | TCAAGCTCAGAGGCCATCTCTTCTCTGACGCTGACCCAAATTAGTCCCAAAATTA | 24 | | | | | |
| Qy | 1563 | ACTGAGAG 1571 | | | | | | |
| Db | 23 | ACTGAGAG 15 | | | | | | |
| RESULT 9 | BE898385 | 677 bp | linear | EST 29-SEP-2000 | | | | |
| LOCUS | 60168121P1 | NIH_MGC_9 | Homo sapiens | CDNA clone IMAGE:3951365 5', | | | | |
| DEFINITION | NIH-MGC | http://mgc.nci.nih.gov/ | | | | | | |
| ACCESSION | BE898385 | GI:10364809 | | | | | | |
| VERSION | BE898385.1 | | | | | | | |
| KEYWORDS | EST. | | | | | | | |
| SOURCE | Homo sapiens (human) | | | | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | | | |
| REFERENCE | 1 (bases 1 to 677) | | | | | | | |
| AUTHORS | NIH-MGC | http://mgc.nci.nih.gov/ | | | | | | |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) | | | | | | | |
| JOURNAL | Unpublished (1999) | | | | | | | |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: DCTD/DTF CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: image.lmnl.gov Plate: L10M819 row: f column: 06 High quality sequence stop: 677. Location/Qualifiers | | | | | | | |
| FEATURES | 1..677 | | | | | | | |
| source | /organism="Homo sapiens" | | | | | | | |
| | /mol_type="mRNA" | | | | | | | |
| | /db_xref="taxon:9606" | | | | | | | |
| | /clone="IMAGE:3951365" | | | | | | | |
| | /tissue_type="adenocarcinoma cell line" | | | | | | | |
| | /lab_host="DH10B (phage-resistant)" | | | | | | | |
| | /clone_id="NIH_MGC_9" | | | | | | | |
| | /note="Organ: ovary; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." | | | | | | | |
| ORIGIN | | | | | | | | |
| Query Match | 24.4% | Score 383.6 | DB 2 | Length 677 | | | | |
| Best Local Similarity | 98.5% | Pred. No. 7.8e-50 | | | | | | |
| Matches 397 | Conservative | 0 | Mismatches | 5 | Indels | 1 | Gaps | 1 |
| Qy | 1090 | CAGGGTATTCCTGGGGGGCTGTGGTCTGCAGATGGCTCCCTGACAGGACTGTGCTCGG | 1149 | | | | | |
| Db | 276 | CAGGGTATTCCTGGGGGGCTGTGGTCTGCAGATGGCTCCCTGACAGGACTGTGCTCGG | 335 | | | | | |

QY 1150 GGAGATTACCTTGTGCGCCGCCAAGACAGACCGGGGTGTCTACAGAACTCTGCAATTC 1209
 |||||
 Db 336 GGAGATTACCTTGTGCGCCGCCAAGACAGACCGGGGTGTCTACAGAACTCTGCAATTC 395
 |||||
 QY 1210 ACCAAGTGATTCAGAGAAACATTCAGGCCCAATCTCTGAGTCATCCAGAGACTCAGACA 1269
 |||||
 Db 396 ACCAAGTGATTCAGAGAAACATTCAGGCCCAATCTCTGAGTCATCCAGAGACTCAGACA 455
 |||||
 QY 1270 CCGGATATCCCGACCTGTGTGAGAGAGACGCCCTGACATCTCTTTCAGACCTCATCTCTTC 1329
 |||||
 Db 456 CCGGATATCCCGACCTGTGTGAGAGAGACGCCCTGACATCTCTTTCAGACCTCATCTCTTC 515
 |||||
 QY 1330 CCAGAGATGTGAGATGTTCATCTCTCAGACCCCTGACCCATGTCTCTGAGACTGAG 1389
 |||||
 Db 516 CCAGAGATGTGAGATGTTCATCTCTCAGACCCCTGACCCATGTCTCTGAGACTGAG 575
 |||||
 QY 1390 GTCTGCTTCCCGACATTTGGGCTGACCGGTCTCTCTAGTTGAAACCTTGGGAAATTTTC 1449
 |||||
 Db 576 GATCTGTTCCTCCCAATTTGGGCTGACCGGTCTCTCTAGTTGAAACCTTGGGAAATTTTC 635
 |||||
 QY 1450 CAAACTGTCTCAGAGCGGGGGTGTCTCTCATCTCTCTGAGG 1492
 |||||
 Db 636 CAAACTGTCTCAGAGCGGGGGTGTCTCTCATCTCTCTGAGG 677
 |||||

RESULT 10 453 bp mRNA linear EST 29-MAR-1996
 N80762/c 298106.g1 Soares fetal lung_NbH19w Homo sapiens cDNA clone
 LOCUS IMAGE:300611.3', mRNA sequence.
 DEFINITION N80762
 VERSION N80762.1 GI:1243463
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 453)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The Washu-Merk EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

FEATURES
 source
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: m13 -40 forward
 High-quality sequence stop: 319.
 Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1245535"
 /db_xref="taxon:9606"
 /clone="IMAGE:300611"
 /dev_page="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_1db="Soares fetal lung_NbH19w"
 /note="Organ: lung; Vector: pTZ19 (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not 1 - oligo(dT) primer [5'-TGTTACCAATCTGATGTGAGAGCGCCGCAATTTTCTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not 1 and cloned into the Not 1 and Eco RI sites of a modified pTZ19 vector (Pharmacia). Library went through one round of

ORIGIN

normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbH19w."

Query Match 24.3%; Score 381.8; DB 7; Length 453;
 Best Local Similarity 96.6%; Pred. 1.7e-49;
 Matches 422; Conservative 0; Mismatches 9; Indels 6; Gaps 3;

QY 1141 GTGTCTGGGAGATTACCTTGTGCGCCGCCAAGAGAAA---CCGGGTGTCTACAGAA 1196
 |||||
 Db 441 GTGTCTGGGAGATTACCTTGTGCGCCGCCAAGAGAAA---CCGGGTGTCTACAGAA 382
 |||||
 QY 1197 CCTCTGCAAGTTACC-AGTGTATCCAGAGAAACCATTCAGGCCAATCTCTGAGTATCC 1255
 |||||
 Db 381 CTCTCTCAAGTTACCAAGATGATCAGAGAAACATTCAGGCCAATCTCTGAGTATCC 322
 |||||
 QY 1256 CAGGACTCAGACACCGGATCCCACTGAC--TGACAGGACAGACCTGACATCTTTTCA 1314
 |||||
 Db 321 CAGGACTCAGACACCGGATCCCACTGAC--TGACAGGACAGACCTGACATCTTTTCA 262
 |||||
 QY 1315 GACCTCATTTCTTCCAGAGATGTGAGATGTTCATCTCTCAGACCCCTGACCCCATG 1374
 |||||
 Db 261 GACCTCATTTCTTCCAGAGATGTGAGATGTTCATCTCTCAGACCCCTGACCCCATG 202
 |||||
 QY 1375 TCTCTGAGTCAAGGTCTGCTTCCCGACATTTGGGCTGACCTGTCTCTAGTTGAAC 1434
 |||||
 Db 201 TCTCTGAGTCAAGGTCTGCTTCCCGACATTTGGGCTGACCTGTCTCTAGTTGAAC 142
 |||||
 QY 1435 CTTGGGAAACAATTTCCAAAATGTCAGAGCGGGGGTGGCTCATCTCTCTGAGGAGCA 1494
 |||||
 Db 141 CTTGGGAAACAATTTCCAAAATGTCAGAGCGGGGGTGGCTCATCTCTCTGAGGAGCA 82
 |||||
 QY 1495 CTTTCATCTCAAGCTCAGAGCCCATCTCTCTGACCTGACCCAAATTTAGTCCC 1554
 |||||
 Db 81 CTTTCATCTCAAGCTCAGAGCCCATCTCTCTGACCTGACCCAAATTTAGTCCC 22
 |||||
 QY 1555 AGAATTAACCTGAGAG 1571
 |||||
 Db 21 AGAATTAACCTGAGAG 5
 |||||

RESULT 11
 AM105502/c 363 bp mRNA linear EST 20-OCT-1999
 LOCUS xds3906.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2597530.3',
 DEFINITION AM105502
 VERSION AM105502.1 GI:6076237
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 363)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.llnl.gov/birp/image/image.html

Possible reversed clone: polyT not found
 Seq primer: -40UP from Gbco

FEATURES High quality sequence stop: 362.
Location/Qualifiers
source 1..363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2597530"
/issue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_11b="NCL CGAP Ov23"
/note="Organ: Ovary; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"

ORIGIN

Query Match 22.6%; Score 355.8; DB 2; Length 363;
Best Local Similarity 99.4%; Pred. No. 1.9e-45;
Matches 357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1213 AAGTGATCCAGAAACCATCCAGGCCAATCTGTGATCATCCAGACTCAGACACCG 1272
DB 363 AAGTGATCCAGAAACCATCCAGGCCAATCTGTGATCATCCAGACTCAGACACCG 304
QY 1273 GCATCCCACTGCTCAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTCCCA 1332
DB 303 GCATCCCACTGCTCAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTCCCA 244
QY 1333 GAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCGACTCAGGGTC 1392
DB 243 GAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCGACTCAGGGTC 184
QY 1393 TGCTTCCCCCAATGGGCTGACCGGTCTCTTAATTGAACCTTGGAACAATTTCGA 1452
DB 183 TGCTTCCCCCAATGGGCTGACCGGTCTCTTAATTGAACCTTGGAACAATTTCGA 124
QY 1453 AACTGTCCAGGAGGGGGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGCTCA 1512
DB 123 AACTGTCCAGGAGGGGGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGCTCA 64
QY 1513 GGGCCCATCTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAAGTGAAG 1571
DB 63 GGGCCCATCTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAAGTGAAG 5

RESULT 12
BM840511 590 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0117576 S12SNU216 Homo sapiens cDNA clone S12SNU216-39-G08 5',
DEFINITION mRNA sequence.
ACCESSION BM840511
VERSION BM840511.1 GI:19196920
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: Yongsung@mail.kribb.re.kr
Plate: 39 row: G column: 08

FEATURES High quality sequence stop: 590.
Location/Qualifiers
source 1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-39-G08"
/sex="F"
/issue_type="Lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"
/clone_11b="S12SNU216"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deapped with tobacco acid pyrophosphatase (TAP). The deapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoR I which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 22.2%; Score 348; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGCTGTGCTGCAATGGCTCCCTCAGGAACTGCTGCTCG 1149
DB 243 CAGGGTATCTGGGGGCTGTGCTGCAATGGCTCCCTCAGGAACTGCTGCTCG 302
QY 1150 GGAAGTTACCCCTGTCGCCGCGCCAAAGACCGGGGTCATACAGGAACCTGCAAGTTC 1209
DB 303 GGAAGTTACCCCTGTCGCCGCGCCAAAGACCGGGGTCATACAGGAACCTGCAAGTTC 362
QY 1210 ACCAAGTGAATCAGAAACCATCCAGGCACTCTGATGATCCAGACTCAGACACA 1269
DB 363 ACCAAGTGAATCAGAAACCATCCAGGCAACTCTGATGATCCAGACTCAGACACA 422
QY 1270 CCGGCATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTCAAGACCTTATCTTC 1329
DB 423 CCGGCATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTCAAGACCTTATCTTC 482
QY 1330 CCAGAGATGTGAGATGTTCACTCTCAGCCCTGACCCCATGATCTCTGAGCTCAGG 1389
DB 483 CCAGAGATGTGAGATGTTCACTCTCAGCCCTGACCCCATGATCTCTGAGCTCAGG 542
QY 1390 GTCTGCTTCCCACTGAGCTGACCGGTCTCTTAATTGAACCT 1437
DB 543 GTCTGCTTCCCACTGAGCTGACCGGTCTCTTAATTGAACCT 590

RESULT 13
BM838406 577 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5',
DEFINITION mRNA sequence.
ACCESSION BM838406
VERSION BM838406.1 GI:19194815
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 577)

AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Beon-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongseung@mail.kribb.re.kr
 Plate: 36 row: A column: 03
 High quality sequence stop: 577.

FEATURES**source**

1. 577
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /clone="S12SNU216-36-A03"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10P"
 /clone_11b="S12SNU216"
 /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 21.2%; Score 333.4; DB 4; Length 577;
Best Local Similarity 99.7%; Pred. No. 4.9e-42;
Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1090 CAGGGGATCTGGGGGGCGTGTGCAATGGCTCCGCGAGGGAGCTGGTCTCTG 1149
 |||||
 243 CAGGGGATCTGGGGGGCGTGTGCAATGGCTCCGCGAGGGAGCTGGTCTCTG 302
 |||||
 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
 |||||
 303 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 362
 |||||
 1210 ACCAAGTGAATCCAGAAACCATCAGGCCCAACTCTGAATGATCCAGAGACTCAGACA 1269
 |||||
 363 ACCAAGTGAATCCAGAAACCATCAGGCCCAACTCTGAATGATCCAGAGACTCAGACA 422
 |||||
 1270 CCGGCAATCCCACTGCTGAGGGAGAGCCCTGACATCTCTTCAGACCTCAATTCCTTC 1329
 |||||
 423 CCGGCAATCCCACTGCTGAGGGAGAGCCCTGACATCTCTTCAGACCTCAATTCCTTC 482
 |||||
 1330 CCAGAGATGTTGAAGATGTTCAATCTCTCAGACCCCTGACCCCAATGTTCTCTGAGACTGAG 1389
 |||||
 483 CCAGAGATGTTGAAGATGTTCAATCTCTCAGACCCCTGACCCCAATGTTCTCTGAGACTGAG 542
 |||||
 1390 GTCTGCTTCCCAATGAGGCTGAGACCGTGTCTCT 1424
 |||||
 543 GTCTGCTTCCCAATGAGGCTGAGACCGTGTCTCT 577
 |||||

RESULT 14
LOCUS W68361/c
DEFINITION 430 bp rRNA linear EST 15-OCT-1996
 Z634008.s1 Soares fetal heart NBH19W Homo sapiens cDNA clone
IMAGE:342591 3', mRNA sequence.

ACCESSION W68361
VERSION W68361.1 GI:1377251
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 430)
 Hillier, L., Clark, N., Dubugue, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST Project

TITLE Unpublished (1995)
JOURNAL Contact: Wilson RK
COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the IMAG Consortium (info@image.lnl.gov) for further information.
 Insert Length: 400 Std Error: 0.00
 Seq primer: mob.RECA-ET
 High quality sequence stop: 286.

FEATURES**source**

1. 430
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="GDB:1267966"
 /db_xref="taxon:9606"
 /clone="IMAGE:342591"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_11b="Soares fetal heart NBH19W"
 /note="Organ: heart; Vector: pT773D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer (5' TGTTCACATCTGAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed from the same fetus as the fetal lung library, Soares fetal lung NBH19W."

ORIGIN

Query Match 20.5%; Score 322.8; DB 7; Length 430;
Best Local Similarity 98.9%; Pred. No. 2.4e-40;
Matches 346; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

1224 GGAACCATCCAGGCGCACTCTGAGTCAAT-CCAGAGACTCAGACACGGGAGATCCCGAC 1282
 |||||
 359 GGAACCATCCAGGCGCACTCTGAGTCAATCCAGAGACTCAGACACGGGAGATCCCGAC 300
 |||||
 1283 CTGCTCAGGAGACAGCGCTGACACTCTTTCAGAC-CACTCTCTCCAGAGATGTTG 1341
 |||||
 299 CTGCTCAGGAGAGACAGCGCTGACACTCTTTCAGAGACCTCTCTCCAGAGATGTTG 240
 |||||
 1342 AGAATGTTCAATCTCTCAGGCGCTGACCCCAATGTTCTCTGAGACTCAGGAGTCTGCTTCCC 1401
 |||||
 239 AGAATGTTCAATCTCTCAGGCGCTGACCCCAATGTTCTCTGAGACTCAGGAGTCTGCTTCCC 180
 |||||
 1402 CACATTGGGCTGACCGTGTCTCTAGTTGAACCTTGGGAGCAATTTCCAAAACTGTCCA 1461
 |||||

Db 179 CACATTTGGGCTGACCGCTGCTCTCTAGTTGAACCTCGGGAACATTTCCAAACCTGTCCA 120
QY 1462 GGGCGGGGGTTGGTGTCTCAATCTCCCTGGGGCACTTTCATCTCCAGCTCAGGGCCCATC 1521
Db 119 GGGCGGGGGTTGGTGTCTCAATCTCCCTGGGGCACTTTCATCTCCAGCTCAGGGCCCATC 60
QY 1522 CTTTCTCTGACGCTGACCCCAATTGATGTCGAGAAATTAAGTGAAG 1571
Db 59 CTTTCTCTGACGCTGACCCCAATTGATGTCGAGAAATTAAGTGAAG 10

RESULT 15
W68496 388 bp mRNA linear EST 15-OCT-1996
LOCUS 234408 r1 Soares fetal_heart_NbHL19W Homo sapiens cDNA clone
DEFINITION IMAGE:342591.5, mRNA sequence.
ACCESSION W68496
VERSION W68496.1 GI:1377366
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 388)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holm, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsch, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Woldmann, P. and
Wilbourn, R.
TITLE The Washu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Washington University School of Medicine
Contact: Wilson RK
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 400 Std Error: 0.00
Seq primer: mob.REGA+RT
High quality sequence stop: 256.
Location/Qualifiers
1.388
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1267966"
/db_xref="taxon:9606"
/clone="IMAGE:342591"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_heart_NbHL19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCATCTGAACTGGAGCGCGCGCATCTTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M. Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."

ORIGIN
Query Match 20.2%; Score 317; DB 7; Length 388;
Best Local Similarity 99.4%; Pred. No. 1.9e-39;
Matches 339; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1224 GGAAACCATCCAGGCCAAGCTCTGAGTCATCCAGAGCTCAGCACACCGGCATCCACC 1283

Db 1 GGAAACCATCCAGGCCAAGCTCTGAGTCATCCAGAGCTCAGCACACCGGCATCCACC 60
QY 1284 TCTGTCAGGAGACAGCCCTGACATCTGCTTTCAGACCTCATTCCTTCCAGAGATGTGAG 1343
Db 61 TCTGTCAGGAGACAGCCCTGACATCTGCTTTCAGACCTCATTCCTTCCAGAGATGTGAG 120
QY 1344 AATGTTCACTCTCCAGCCCTGACCCCAATGTCCTGAGACTCAGAGTCTGCTCCCA 1403
Db 121 AATGTTCACTCTCCAGCCCTGACCCCAATGTCCTGAGACTCAGAGTCTGCTCCCA 180
QY 1404 CATGGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTTCCAAAATGTCAGG 1463
Db 181 CATGGGCTGACCGTGTCTCTAGTTGAACCTGGGAAACAATTTCCAAAATGTCAGG 240
QY 1464 GGGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTCATCTCAGCTCAGGGCCCATC 1522
Db 241 GGGGGGGTGGCTCTCAATCTCCCTGGGGCACTTTCATCTCAGCTCAGGGCCCATC 300
QY 1523 CTTTCTGTC-AGCTGACCCCAATTGATGTCGAGAAATTA 1562
Db 301 CTTTCTGTCAGCTGACCCCAATTGATGTCGAGAAATTA 341

Search completed: February 27, 2005, 12:42:30
Job time : 4932.95 secs

|||||
Db 301 ACATCTGTGCACACAGGCACTGTGTGGCACTGAGACATGTGGATGAGGTGTGGC 360
QY 3116 ATCCCCCTAGGCTGCCCGGAGGCTGTGTACTCTGAGACAGAGCTGTATGTATGTGTCAC 3175
Db 361 ATCCCGCTAGGCTGCCCGGAGGCTGTGTACTCTGAGACAGAGCTGTATGTATGTGTCAC 420
QY 3176 CTGTGAGGCAATGAGGCTGTGTGTGAGAACTGCGGTGCTGTGCTGTATCTGTGT 3235
Db 421 CTGTGAGGCAATGAGGCTGTGTGTGAGAACTGCGGTGCTGTGCTGTATCTGTGT 480
QY 3236 TGTGCGGTGTCTTGTGGGTGAGTTCTGTAATGATGTGTGTGCAAGGCAATCAGCAAG 3295
Db 481 TGTGCGGTGTCTTGTGGGTGAGTTCTGTAATGATGTGTGTGCAAGGCAATCAGCAAG 540
QY 3296 GGTAAAGAACAGGCGCGGCGGTGTGTCTGCTGTAAATCCAGGCTTTTGGAGGCGC 3355
Db 541 GGTAAAGAACAGGCGCGGCGGTGTGTCTGCTGTAAATCCAGGCTTTTGGAGGCGC 600
QY 3356 AGGCAAGCGGATCACTGAGGTGCGGAGATCGAGGCGAGCTGACCAATGAGAAACC 3415
Db 601 AGGCAAGCGGATCACTGAGGTGCGGAGATCGAGGCGAGCTGACCAATGAGAAACC 660
QY 3416 CGTCTCTACTAAATAATCAAAAAATTAGCTGTGTGTGTGCGCGCTGTGTAAATCCAGC 3475
Db 661 CGTCTCTACTAAATAATCAAAAAATTAGCTGTGTGTGTGCGCGCTGTGTAAATCCAGC 720
QY 3476 TACTCGGGAGACTGGGGCAGAAAAATCGCTGTAAACCCGGAAGTGAAGTTGGGTGAGC 3535
Db 721 TACTCGGGAGACTGGGGCAGAAAAATCGCTGTAAACCCGGAAGTGAAGTTGGGTGAGC 780
QY 3536 CGAGATCGCGCACTTCGACTCCAGCTGGGCAACAAGAGGAAATCCGTCTGAAAGAA 3595
Db 781 CGAGATCGCGCACTTCGACTCCAGCTGGGCAACAAGAGGAAATCCGTCTGAAAGAA 840
QY 3596 AAAAAAGAAAAAAGGTAAGAACCAATGTAATGGGCAAGGAGCACTGATATGAGT 3655
Db 841 AAAAAAGAAAAAAGGTAAGAACCAATGTAATGGGCAAGGAGCACTGATATGAGT 900
QY 3656 GGGGCACTGCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 3715
Db 901 GGGGCACTGCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 960
QY 3716 GGCATGTTTCACTGTGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGGCC 3775
Db 961 GGCATGTTTCACTGTGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGGCC 1020
QY 3776 CTAACTGAGCCCTTCTTCTGTGTCTGTCTTCTGAAACCTAGCTTCCGCTCATGAGCTC 3835
Db 1021 CTAACTGAGCCCTTCTTCTGTGTCTGTCTTCTGAAACCTAGCTTCCGCTCATGAGCTC 1080
QY 3836 TGACCCCACTCTTCTTCTTCTTCAACAGCCCTTGGGCAAGCTTATGTGACCCCCCTTAA 3895
Db 1081 TGACCCCACTCTTCTTCTTCTTCAACAGCCCTTGGGCAAGCTTATGTGACCCCCCTTAA 1140
QY 3896 GGCACACCCCTTGTGGGCAAGCTTCAACCCCTATTCTGTGGGTACCTTTTAAACCCCC 3955
Db 1141 GGCACACCCCTTGTGGGCAAGCTTCAACCCCTATTCTGTGGGTACCTTTTAAACCCCC 1200
QY 3956 TTCAAAAGTCAGAGCTTTTCTTTTCTTTTGTGAGACAGTCTTGTCTCTCCAGGC 4015
Db 1201 TTCAAAAGTCAGAGCTTTTCTTTTCTTTTGTGAGACAGTCTTGTCTCTCTCCAGGC 1260
QY 4016 TGAAGTGCAGTGTGCGGATCTGCGCTCACTGCAACCTCTTGTCCCAAGTTCAAGTATT 4075
Db 1261 TGAAGTGCAGTGTGCGGATCTGCGCTCACTGCAACCTCTTGTCCCAAGTTCAAGTATT 1320
QY 4076 CTGTGCTCTCACTCTCTGAGTGTGAGTTACAGGTGTGCGGCAACAGCTTGTGCTAA 4135
Db 1321 CTGTGCTCTCACTCTCTGAGTGTGAGTTACAGGTGTGCGGCAACAGCTTGTGCTAA 1380
QY 4136 TTTTGTGTCTTGTAGTGAAGACAGGTTTCACTTGTGTGCGAGGCTGTGCTCAAACTCC 4195
|||||

Db 1381 TTTTGTGTCTTGTAGTGAAGACAGGCTTTCACTTGTGTGGCAAGCTGTCTCAAACTCC 1440
QY 4196 CAACCTCAGGTATCCGCCCACTGGGCTCCCAAGGTGCTGGGTTACAGGCGTGAAGCC 4255
Db 1441 CAACCTCAGGTATCCGCCCACTGGGCTCCCAAGGTGCTGGGTTACAGGCGTGAAGCC 1500
QY 4256 ACCGCCCCAGGCCCAAGTCAAGGCTCTTATAGAGACTCTTAACATGTAACTTGAACC 4315
Db 1501 ACCGCCCCAGGCCCAAGTCAAGGCTCTTATAGAGACTCTTAACATGTAACTTGAACC 1560
QY 4316 TGGCTTAACTAATGTAATTCAAACCCCTTCTGTCTGCTGACGCTTCACTCACTG 4375
Db 1561 TGGCTTAACTAATGTAATTCAAACCCCTTCTGTCTGCTGACGCTTCACTCACTG 1620
QY 4376 AGGCTGAGCCCACTTCTTGAACCAAGTTCATCCCTTAAAGGCTGTGTCTCTCCATC 4435
Db 1621 AGGCTGAGCCCACTTCTTGAACCAAGTTCATCCCTTAAAGGCTGTGTCTCTCCATC 1680
QY 4436 CCCAGGCTCAGACCCCAAGCTTGGCACTACCCCTGAGCTTGTTCAGAAATCCGTGAC 4495
Db 1681 CCCAGGCTCAGACCCCAAGCTTGGCACTACCCCTGAGCTTGTTCAGAAATCCGTGAC 1740
QY 4496 CCAATTTTACCTTCACTATGATTTCAAGCTTCAAGTTCAGAAATCTGTGAGTCAATGAC 4555
Db 1741 CCAATTTTACCTTCACTATGATTTCAAGCTTCAAGTTCAGAAATCTGTGAGTCAATGAC 1800
QY 4556 GTTCAAGTACCTTCACTGAGCTTGGGCTGTGCTTGAAGCTTGAAGCTTGGGCTTGAAG 4615
Db 1801 GTTCAAGTACCTTCACTGAGCTTGGGCTGTGCTTGAAGCTTGAAGCTTGGGCTTGAAG 1860
QY 4616 TGCATCTTATTTCTCAGAGCCCTGCGCTGCGCTCAGCATGTCAAGACCCCTTC 4675
Db 1861 TGCATCTTATTTCTCAGAGCCCTGCGCTGCGCTCAGCATGTCAAGACCCCTTC 1920
QY 4676 TAGCTGTGTGCTCTTGTGATGTGAACCCACCCCAAGCCCGCTTGTGAGCC 4735
Db 1921 TAGCTGTGTGCTCTTGTGATGTGAACCCACCCCAAGCCCGCTTGTGAGCC 1980
QY 4736 CGGCCCAACCTTCTTCTGCTCCAGAGATGTCTGCGCAATGATGTTCCGTGTA 4795
Db 1981 CGGCCCAACCTTCTTCTGCTCCAGAGATGTCTGCGCAATGATGTTCCGTGTA 2040
QY 4796 CCACCTCTTAAACACCTGCTCTTGGAGCAACAGGACTTGGAGCTGGGCGGAG 4855
Db 2041 CCACCTCTTAAACACCTGCTCTTGGAGCAACAGGACTTGGAGCTGGGCGGAG 2100
QY 4856 AAGCGCCGCTGTGATGACAGCAAGCCGATCAATCAATGATTCGATTCGATATGCA 4915
Db 2101 AAGCGCCGCTGTGATGACAGCAAGCCGATCAATCAATGATTCGATTCGATATGCA 2160
QY 4916 CACCCAGCGGTGGAGGCGCGCTGTGCTAAGGCGCAACCAAGCTTCTGCGGCGGT 4975
Db 2161 CACCCAGCGGTGGAGGCGCGCTGTGCTAAGGCGCAACCAAGCTTCTGCGGCGGT 2220
QY 4976 GTTGTGATTCACAGTGTGCTCTC 5000
Db 2221 GTTGTGATTCACAGTGTGCTCTC 2245

RESULT 2
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommel, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereeto

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


|||||
Db 17544 TTATAGAGAGACGGGTTTACCGTGTATGACGAGATGCTTGAATCTCTGACCTCGTG 17603
597 ATCCGCCCGCCCTCGGCGCTCCCAAAGTGTGGGATTACAGCCGTGACCCACCGGCCCGGCG 656
Db 17604 ATCCACTCTGCTCAGCTCCCAAAGTGTGGGATTACAGCCGTGACCCACCGGCCCGTAC 17663
Qy 657 CA-----TGAATCTCTTGAATGCTATGCTATGTA-----CAAGTACTTAAGCCA 703
Db 17664 CAATTTATTTTCACTTGTGCTGGGCTTTTATTAATTAATTAATTAATTAATTAATTA 17723
Qy 704 TCAGACTTACCTTTAATATGATGATTTGGGCGAGGCACTGCTCATGCTGTAAAT 763
Db 17724 CACAAAGATGATATAGATATGATGAGGCGAGGTGACGAGGCTCACACTGTAAATC 17783
Qy 764 CCAGCACTTTGGGAGGCGAGGTTGGGATGATCTTGAAGCCGAGGCTTGGAGACAGCC 823
Db 17784 CTAGCACTTTGGGAGGCTGAGGTTGGGATGATGCTGAGGCTTGAAGGCTTGAAGCC 17843
Qy 824 TGGCCCAATGTGTAATCTGTCTTACTTAAATTAATTAATTAATTAATTAATTAATTAAT 883
Db 17844 TGGCCCAATGTGTAATCTGTCTTACTTAAATTAATTAATTAATTAATTAATTAATTA 17892
Qy 884 CGGCTGTGTGGGCGACACCTGTAAATCCAGCTATGCTGAGGCTGAGGCGAGGATCA 943
Db 17893 TGGATGTGTGGGCGACACCTGTAAATCTGACTACTGAGGAGGCTGAGGCGAGGATCA 17952
Qy 944 CTGGAACCTTGGGAGGCGGAGTTGCACTGGGCGGAGATCAATCAATCAATCAATCAATCA 1003
Db 17953 CTGGAACCTTGGGAGGCGGAGTTGCACTGGGCGGAGATCAATCAATCAATCAATCAAT 18012
Qy 1004 GGGCAGAGAGCAAGACTGTGTCTCAATTAATTAAT 1039
Db 18013 GGGCAGAGAGCAAGACTGTGTCTCAATTAATTAATTAAT 18048

RESULT 7
US-09-949-016-15787/C
; Sequence 15787, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15787
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15787

Query Match 8.4%; Score 417.6; DB 4; Length 120727;
Best local Similarity 80.5%; Pred. No. 5.8e-87;
Matches 546; Conservative 0; Mismatches 109; Indels 23; Gaps 4;

Qy 362 TTTTGT 421
Db 32483 TTTTGT 32424

Qy 422 ATTCGGCTCACTGCAAGCTCCGCTCCCGGGTTGACAGCATTTCTCTGCTCAGGCTCC 481
Db 32423 ATTCGGCTCACTGCAAGCTCCGCTCCCGGGTTGACAGCATTTCTCTGCTCAGGCTCC 32364
Qy 482 CAATGCTGTGGAATCAAGGCGCCCGGCTCACTACGCCGCTAA-TTTTTTGTATTTTAA 540
Db 32363 CAATGCTGTGGAATCAAGGCGCCCGGCTCACTACGCCGCTAA-TTTTTTGTATTTTAA 32304
Qy 541 GTTAGAGCGGGTTTCAACGTTTGAAGCGGAGTGGCTGATCTCTGACCTCGTATCC 600
Db 32303 GTTAGAGCGGGTTTCAACGTTTGAAGCGGAGTGGCTGATCTCTGACCTCGTATCC 32244
Qy 601 GCCCGCTTCGCTCCCAAAGTGTGGAATTAACAGCGTGAAGCCAACCGGCCGCGCATG 660
Db 32243 ACCCACTGCGCTCCCAAAGTGTGGAATTAACAGCGTGAAGCCAACCGGCCGCGCATG 32184
Qy 661 ATCATCTTTTACATGCTATGCTATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 720
Db 32183 ATCATCTTTTACATGCTATGCTATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 32127
Qy 721 AATATGAGTTTGGGCGAGGCGACCGGCTCATGCTGTAAATTCAGCACTTTGGAGGC 780
Db 32126 -----AGACTTGGCTGTGACGGTGGCTCACACTGTGTATCTTACACTTTGGAGGC 32074
Qy 781 AGAGGTGGGTGAATCACTTGAAGCGGAGGTTTGAAGCAAGCTGAGCCAAATGTGAAA 840
Db 32073 AGAGGTGGGTGAATCACTTGAAGCGGAGGTTTGAAGCAAGCTGAGCCAAATGTGAAA 32014
Qy 841 CTCTGCTTTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Db 32013 CTCTGCTTTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 31966
Qy 901 ACCCTGAATCCAGCTATGCTGAGGCTGAGGCGAGGAGTCACTTGAACCTTGGAGGC 960
Db 31965 ACCCTGAATCCAGCTATGCTGAGGCTGAGGCGAGGAGTCACTTGAACCTTGGAGGC 31906
Qy 961 GAGGTGCAAGTGGCGGAGATCAATCAACCGCTTCAGGCTGGGCGAGAGCAAGACT 1020
Db 31906 GAGGTGCAAGTGGCGGAGATGCGGCTTCAGCTTCAGGCTGGGCGAGAGCAAGACT 31846
Qy 1021 CTGTCTCAATTAATTAATTAAT 1038
Db 31845 CTGTCTCAATTAATTAATTAAT 31828

RESULT 8
US-09-949-016-15788/C
; Sequence 15788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15788
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15788

```
; SEQ ID NO 16038
; LENGTH: 166698
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (166698)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16038

Query Match      8.3%; Score 415.4; DB 4; Length 166698;
Best Local Similarity 76.2%; Pred.No. 2.2e-86;
Matches 560; Conservative 0; Mismatches 151; Indels 24; Gaps 3;

OY       328 TGGGTCTATGATCATCTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGACGAGAGTC 387
DB       106710 TAGCTTGTTTAATTATAATTAAATTAAATTAAATTAAATTAAATTAAATTGAAGTAGGTCT 1067658

OY       388 CGCTCTGCGCCCAAGCGCTGAGTGCAGTGGCGGAGTCTCGACTCAGCTCCAGCTCCGCT 447
DB       106770 TGCTCTGCGCCCAAGCGCTGAGTGCAGTGGCGAGTCTCTGCTCACCGCAGCTCCGCT 1068282

OY       448 CCCGGGTTCAAGCCATTCTCTCTGCTCAGCCCTCCCAAGTAGCTGGGACTACAAGCCGCC 507
DB       106830 CTTGGGTTCAAGCCATTCTCTCTGCTCAGCCCTCCGATTAGCTGGGATTACAGGCCCCG 1068858

OY       508 CCACTAGCCCCGGCTAATTTTTTTTGTATTTTGTATAGAGACCGGGTTTCAACGTTTAC 567
DB       106890 CCACCAGCCCAGATTAATTTTTTTTGTATTTAGTAGAGTGGGTTTCACTGTGTAGC 1069484

OY       568 CGGATGCGCCCGATCTCTGTAACCTCGTGATCCGCCCGCTCGGCTCCCAAAGTGTCTG 627
DB       106950 CAGATGCTCTCAATCTCTCTGTAACCTCGTGATCCGCCCTCTCGGCTCCCAAAGTGTCTG 1070099

OY       628 GATTACAGGCGCTGAGCCACCGCGCCCGGCAATGATCATCTTCTTGACTATGCTGATGTGA 687
DB       107010 GATTACAGGCGCTGAGCCACCGCGCCCGGCACTTGTAAAAAACAAAACATATTACTG 1070658

OY       688 CAAGTACTTAAAGCC-----ATCAGA CTCTACCCCTTAAATATGAGTTGGGC 736
DB       107070 AAAAATAAATGTGCCATGAGATTGTGAAGAATTAGCATTAATAATAAGGCG 1071292

OY       737 CAGGCA-CCTGTGGCTCATGCTCTGTAA TTCAGCACTTTGGAGGCGAGAGCTGGGTAA TC 795
DB       107130 TTТААAGTGTGGCTCAGCGCTGTAA TC CAGCAATTTTGAAGGCTGAA CCGGSCGATC 1071898

OY       796 ACTTAGGCGCAGAGTTTGAGACCGACGCTGGCCAATGSTRAAA CTCGTCTTACTPA 855
DB       107190 ATTГAGGTCAGAGTTTGA AACCGACTGGCCAATGTBTAA ACCCGCTCTACT-- 1072477

OY       856 AAAAAAAAAAAAAAAAAAATTCAGCCGGGTGTCTGTGGGCGACACTGTATCCAG 915
DB       107248 -----AAAAATACAAAATTTAGCAAGCGCTGTGTGGCACACGCTGTATCCAGC 1072977

OY       916 TATGCTGAGGCTGAGGCA CAGAGTCACTTTGA ACCTTGAGGCGGAGGTTGCA GTGGC 975
DB       107298 AACTCAGAGGCTGTGAGCAGAGAGATCGCTTGA ACCAGAAAGTGTGACCTTGCAGTGAGC 1073577

OY       976 CGAGATCAATCACCGCCCTCAGCTGTGGGCGAGAGCAAACTCTGTCTMAAATAAT 1035
DB       107358 TGAATATCTGCATTGTCACTCCAGCTGTGGTGA CAGAGCAAGTCTCATCTCAAAAAA 1074177

OY       1036 AATTAACAAACGAA 1050
DB       107418 AAAAAAAAAAAAAA 107432

RESULT 10
US-09-949-016-16080
; Sequence 16080, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

?
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001107
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 15080
? LENGTH: 79858
?
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-16080

```

| | | | | | | | |
|-----------------------|--------|--------------|--------|------------|------|--------|--------|
| Query Match | 8.3%; | Score | 414.2; | DB | 4, | Length | 79858; |
| Best Local Similarity | 76.2%; | Pred. NO. | 3e-86; | | | | |
| Matches | 557; | Conservative | 0; | Mismatches | 128; | Indels | 46; |
| | | | | | | Gaps | 2; |

| | | | |
|----|-------|---|-------|
| OY | 349 | TTTTTTTTTTTTTTTTTTTTTTTTTTTGAACGAGATCGCTCTGTGCGCCAGGCTGGA | 408 |
| Db | 10498 | TTTTTTTTTTTTTTTTTTTTTTTTTTTGAACGAGATCGCTCTGTGCGCCAGGTTGGA | 10557 |
| OY | 409 | GTGCAGTGGCGGAGTCTCGGCTCACTGCAAGCTCCGCTCCGAGTTCAGCCATCTCC | 468 |
| Db | 10558 | GTGCAATGGGTGATCTTGGCTCACTGCAAGCTCAATCTCCGGTTCATTCATCTCTCC | 10617 |
| OY | 469 | TGCTCAAGCTCCCAAGTAGTGGAGTACAGGCGCCCGCACTACGCGCGCTAATTTT | 528 |
| Db | 10618 | TGCTCAAGCTCCCGAATAGTGGAGTACAGGCGCGCCGCAACAAGCTGGCTAATTTT | 10677 |
| OY | 529 | TTTGTATTTTATAGAGAGCGGGTTTCAAGTTTTAGCCGGAGTGGCCCTGATCTCTG | 588 |
| Db | 10678 | TTTGTATTTTATAGAGAGAGGGTTTCAACATGTTAGCCAGCTGTCTCAATCTCTG | 10737 |
| OY | 589 | ACCTGTGATCCGCGCGCTCGGCTCCAAAGTGTGGATTACAGGCGTGAGCCACCG | 648 |
| Db | 10738 | ACCTGTGATCCGCGCGCTCTGCTCTCCAAAGTGTGGGATTACAGGCGTGAGCCACCG | 10797 |
| OY | 649 | CGCCCGCCAGATCATCTTCTTGACATAGCTGATGCAAT----- | 692 |
| Db | 10798 | CGCCCGCCAGAAATCTTCTTCTACTCTTAACAACAGATCAATAGTCATCTAATATTTTC | 10857 |
| OY | 693 | -----ACCTAAAGCCATCAACTCTACCTTTAATATGCAATTGTTGG | 734 |
| Db | 10858 | TTCTTAAAAAACGTAGATGATNTAATTAATTCATACATAGCTAATTTAATACATACAAAAG | 10917 |
| OY | 735 | GCCAGGCAAGGTGCTCATGCTGTAAATTCACAGACTTTTGGAGGCGAGAGTGGTGAAT | 794 |
| Db | 10918 | GCCAGATGCGGTGCTCATGCTGTAAATTCACACACTTTTGGAGGCGGTGAGT | 10977 |
| OY | 795 | CACCTGAGGCGAGAGTTTGAGACAGCTGGGCGCAATGTGTAAACTGTCTTTACATA | 854 |
| Db | 10978 | CACCTGAGGTCAGAGTTTCAAGACAGCTGGGCGCAAGATGTGTAAATTCGCTTCTACT- | 11036 |
| OY | 855 | AAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTGTTGGGCGACACCTGTAAATCCAG | 914 |
| Db | 11037 | -----AAAAATATCAAAAATTAGCCAGGCGTGTGTGGCGGACCTGTAAATCTCAA | 11085 |
| OY | 915 | CTATGCTGAGAGCTGAGAGGCAAGAGATCACTTGAACCTTGTGAGGCGGAGTTGCAATGGG | 974 |
| Db | 11086 | CTACTGGGGTGACTGAAGGCAAGGAAATCGCTTAAACCCAGAGAGCGAAAGTTTGGTGGAG | 11145 |
| OY | 975 | CCGAGATCATCTACCGCTCCAGAGCTGGGCGCAAGAGCAAGACTGTGTCMAATTAAG | 1034 |
| Db | 11146 | CCAGATCATCGCCACTGCACTCTGCTGGGCGATGAGTGAAGCTCTGCTTCMAAATAA | 11205 |
| OY | 1035 | TAAATTAACAAG | 1045 |

```

DB              11206 AAAAAAAAAA 11216

RESULT 11
US-09-949-016-12218/c
; Sequence 12218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12218
; LENGTH: 60304
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12218

```

| Query Match | 8.3%; | Score 414; | DB 4; | Length 60304; |
|---------------------------|--------|--------------------|------------|---------------|
| Best Local Similarly | 78.5%; | Pred. No. 2.9e-86; | | |
| Matches 551; Conservative | 0; | Mismatches 135; | Indels 16; | Gaps 4 |

| | | | | | | | |
|----|-------|--|---------------------|---------|---------|-------------------------------------|-----|
| OY | 352 | TTTTTTT | TTTTTTT | TTTTTTT | TTTTTTT | TGAGACGAAGTCCTGGCTGTGGCCCAAGCTGAGNG | 411 |
| Dd | 36261 | TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGATGAGAT---- | CTGTGCTGAGGCTGAGATT | 36206 | | | |
| OY | 412 | CAGTGGCGGAGTCCTGGCTCATCTGCAGAGCTCCGCTCCCGGGTTCAGGCATCTTCCTGC | | 471 | | | |
| Dd | 36205 | CAGTGGCGGAACTCTCGGCTCAGCTGCAGAAGCTCCGCTCCGCGTTACAGGCATTTCTCTGC | | 36146 | | | |
| OY | 472 | CTCAGCTCCCAAGTAGCTGGAGCTAACAAGCGCCGCCCATACAGCCCGGCTAATTTTTTT | | 531 | | | |
| Dd | 36145 | CTCAACTCCAGATGTAGCTGGGCTACAGCGGCCCAACACACCGCCGGCTAATTTTTTT | | 36086 | | | |
| OY | 532 | GTAATTTTATAGAGACGGGGTTTACCGTTTAAAGCCGGAATGGGCTCGATCCCTGACC | | 591 | | | |
| Dd | 36085 | GTAATTTTATAGAGACGGGGTTTACCGTTTAAAGCAATGGTCTCAATCTCTGACC | | 36028 | | | |
| OY | 592 | TCTGTATCCGCCCGCTCTGGCTCCCAAAGTGTGGGATTACAGCGTSGACACCGCGC | | 651 | | | |
| Dd | 36025 | TCTGTATCCGCCCTGGCTCAGCCTCCCAAAGTGTGGGATTACAGCGTSGATGAGC-CAC | | 35967 | | | |
| OY | 652 | CCGCCATGATCATCTTCTTGAATAAGCTGATGTGACAAAGTACTTAAAGCCATCAGACTC | | 711 | | | |
| Dd | 35966 | CGTCCCGGCCCGCCCCAGGCTGGTTCTATGCTCTCAAGCTTCTGATCCACTCTGG | | 35907 | | | |
| OY | 712 | TACCTTAAATATGAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAATCCAGACT | | 771 | | | |
| Dd | 35906 | AACCTTGTCAAGCTGAGTTAGCCAGGACAGTGGCTCATACCTAATAATCCAGACT | | 35847 | | | |
| OY | 772 | TTGGAGAGCAGAGTGGGTGAATCACTTGAAGCCAGAGTTTGAGACAGCCTGGCCAAC | | 831 | | | |
| Dd | 35846 | TTGGAGATGTGAAGTGAATGATCACTTGAAGTCAGAGATTCAAGCACCGCTGACCAAC | | 35787 | | | |
| OY | 832 | ATGTGAAATCTGTCTTTTAAATAAAAAAAATAAAAAAAATGACCGGCTGTC | | 891 | | | |
| Dd | 35786 | ATGTGAAATCCATCTTACT-----AAAATATCAAAAAAATTAGCCAGGCTGA | | 35737 | | | |
| OY | 892 | GTGGGGCAGACCTGTATATCCAGACTATGCTGAGGCTGAGGACAGAGATCACTTGAAC | | 951 | | | |
| Dd | 35736 | GTGTGTACACCCCTGTATATCCAGACTTATAGAGGCTGAGGACAGAGATCACTTGAAC | | 35677 | | | |
| OY | 952 | CTGAGGCGGAGTTGACGTGGCCGAGATCATCACCGCCCTTCAGCTGGGCGACAG | | 1011 | | | |

us-09-936-271b-13 copy 1 5000.rnd

| Qy | 892 | GTGGGGACAC | CMGTATATCC | AGCTATGC | TGAGGGGAC | AGAGTCACTT | GAAC | 951 |
|----|-------|---|---|---|---|---------------------------------------|------|-------|
| Qy | 892 | GTGGGGACAC <td>CMGTATATCC <td>AGCTATGC <td>TGAGGGGAC <td>AGAGTCACTT <td>GAAC</td> <td>951</td> </td></td></td></td> | CMGTATATCC <td>AGCTATGC <td>TGAGGGGAC <td>AGAGTCACTT <td>GAAC</td> <td>951</td> </td></td></td> | AGCTATGC <td>TGAGGGGAC <td>AGAGTCACTT <td>GAAC</td> <td>951</td> </td></td> | TGAGGGGAC <td>AGAGTCACTT <td>GAAC</td> <td>951</td> </td> | AGAGTCACTT <td>GAAC</td> <td>951</td> | GAAC | 951 |
| Db | 35736 | GTGGTAAACCC | CMGTATATCC | AGCTCTT | TGAGGGGCTG | AGGACAGAAATCACTT | GAAC | 35677 |
| Qy | 952 | CTGGAGCGGAG | GGTTGAGT | GGGCGAGAT | CAATACAGCCCT | CAGGCTGGGCGACAG | | 1011 |
| Db | 35682 | CTGGAGCGGAG | GGTTGAGT | GGGCGAGAT | CAATACAGCCCT | CAGGCTGGGCGACAG | | 1011 |

```

DB
35617 AGAGGACTGTCTCAAAAACAAAACCAAAAAACCCAA 35676

RESULT 13
US-09-949-016-11980/c
; Sequence 11980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; TYPE: DNA
; ORGANISM: Human

```

```

; LOCATION: (1) ... (125188)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11980

Query Match      8.2%; Score 411.2; DB 4; Length 125188;
Best Local Similarity 78.7%; Pred. No. 1,9e+85;
Matches 535; Conservative 0; Mismatches 128; Indels 17; Gaps 3;

```

2
2
2
2
2
2
2
2

[illegible]

111111

120E

[illegible]

ITCCA 766

Db 120428 GCATAT-----AGAAITCTTAAAGTTCGAGCGGCAATGATGGCTCATCTTCTAATCCCA 12037

| | | | |
|----|--------|--|--------|
| Qy | 767 | CCACCTTTGGAGGACGAGGTGGGTGAATCACTTTAGAGCCAGAGTTTGAACACAGCTGG | 826 |
| Db | 120312 | GCACCTTTGGAGGCGCAGGCGGACAGATACACTTGAAGTTTGAAGTTGAGACCAAGCTGG | 120313 |
| Qy | 827 | CCAACATGCTGAACCTCTGTCTTTACTAATAAAAAAAAAAAAAAAAAAAATCAGCCGG | 886 |
| Db | 120312 | CCAACATGCTGAACCCCTATTCTACT-----AAAAATCAAAAAAATACCTGG | 120265 |
| Qy | 887 | GTGTGTGGGGGACACCTGTATCCAGCTATGCTGAGAGCTGAGGCGACGAGATCACTT | 946 |
| Db | 120264 | GTGGGGTGGCACGCGCTGTATGTCGCCAGTATTTGGGAAGCTTAACTGAAGATACTT | 120205 |
| Qy | 947 | GAACCTTGAGGCGGAGGTTGCAGTGGGCGGAGATCACATCAACCGCCCTTCAGCCTGGGC | 1006 |
| Db | 120204 | GAACCCGGAGGTTGAGGTTGCAGTGAAGCCAAAGATGCAAGCACTGTAAACCCAGCGTGGGT | 120145 |
| Qy | 1007 | GACAGAGCAAGCTCTGTCT | 1026 |
| Db | 120144 | GACAGAGCGAGACTCTGTCT | 120125 |

```

RESULT 14
US-09-949-016-13861
; Sequence 13861, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13861
; LENGTH: 33379
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(33379)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13861

```

| Query Match | 8.1% | Score 405.6 | DB 4 | Length 33379 |
|-----------------------|--------------|---|----------------|--------------|
| Best Local Similarity | 70.0% | Pred. No. 2e-84 | | |
| Matches 633 | Conservative | 0 | Mismatches 209 | Indels 62 |
| | | | Gaps | 4 |
| QY | 335 | TATGATCAATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACGGAGTCCGCTCG | 394 | |
| Db | 5215 | TATGGTGNAAATTAATCTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGCGAGTCCGCTCG | 5274 | |
| QY | 395 | TCGCCAGGCTGGAAGTGCAGTGGCGGAGTCTGGGCTCACTGCAGCTCCGCTCCGGGT | 454 | |
| Db | 5275 | TCGCCAAGCTGGAAGTGCAGTGGCGGAGTCTGGGCTCACTGCAGCTCCGCTCCGGGT | 5334 | |
| QY | 455 | TCAGGCATTTCCCGGCTCAGCCCTCCCAAGAGTGGGACATACAGGGGCCCGGCACATAC | 514 | |
| Db | 5335 | TCAGGCATTTCTCTGCGCTCAGCCCTCCCAAGAGTGGGACATACAGGGGCCCGGCACATAC | 5394 | |
| QY | 515 | GCCCGCATTAATTTTTTTGTAATTTTTTAAGTAGAGACGGGGTTTCAACGTTTAAAGCCGGAGTG | 574 | |
| Db | 5395 | GCCCGCATTA-TTTTTTGTAATTTTTTAAGTAGAGACGGGGTTTCAACGTTTAAAGCCAGATG | 5455 | |
| QY | 575 | GCGTGAATCTCTGACCTCGTGATCCGCCGCTTGCGCTCCCAAAAGTCGAGGATTACA | 634 | |

| | | | |
|----|------|---|------|
| Db | 5454 | GTCGGAATCTCCTGACTCGTAGGCACGCCGGCCCTCAGCCTCCCAAAAGTGTGGATTACA | 5513 |
| OY | 635 | GGCGTGAGCCACCGCGCCCGGCANTGATCATCTTCTT----- | 671 |
| Db | 5514 | GGTGTGAGCCACC GGCGCCGGCCGGTGTAAAATTTATCTTAGCCATGTGTTCAAATGA | 5573 |
| OY | 672 | -----GACTAATGCTGATGTGACAGATGACTTAAGCCATCA | 706 |
| Db | 5574 | ACAAAGTTGTACATGAAGAAGTTTAGTAAATTTTTGTAAATTTAAATATGTATGACTGTAA | 5633 |
| OY | 707 | GACTCTAACCCCTTAAATATGCAAGTTTG3GGCAGGCAAGCCGTGCTCATGCTGTAAATCCA | 766 |
| Db | 5634 | GATATTTAGTGTTTTAAAGAGTTTCAGGGCCGGCCGTGTGTGCTCACGCGCTGATATCCA | 5693 |
| OY | 767 | GCACCTTTGGAGGCGAGAGTGGGTGGTAATCACTTGAGGCGCAGAGTTTGAGACGAGCTCTGG | 826 |
| Db | 5694 | GCACCTTTGGAGGCGCGAGGCGAGGTGATCAC--GAGGTCAAGATATCAAGCCATCTCTGG | 5751 |
| OY | 827 | CCAACATGATGTAACTCTGCTCTTACTTAAAAAATAAAAAAAAAAAAAAAAAATCAGCCGG | 886 |
| Db | 5752 | CTAAACATGATGTAAACCCCATCTCTACT-----AAAAATGCAAAAAAATTAGCCGG | 5800 |
| OY | 887 | GTCGTCGTGGGCGACACTGTATATCCACGCTATGTGTGAGGCTGAGGCGACAGAGTCACTT | 946 |
| Db | 5801 | GTCGTGTGTGCGACGCGCCTTAATATCCACGTACTCTTGAGGTTTGAGGCGAGTAATGTGCAT | 5860 |
| OY | 947 | GAACCCGTGGAGGCGGAAGTTGTGCACTGGGCGAGATCATCAACGCCCTCCAGCCTG3GC | 1006 |
| Db | 5861 | GAACCTTGGAGGCGGAGACTGTGAGTGAAGCGAGATCTGTGCCACTGTCAAGCTTG6GC | 5920 |
| OY | 1007 | GACGAGCCAAAGACTCTGTCTTCAAAATTAATAATTAACAAACGAACGAAGTTGTTGTA | 1066 |
| Db | 5921 | GACGAGCCGAGACTCTGTCTCAAAAAAATAAAAAAAAAAGNAGGTTTCAATAAATG3GCT | 5980 |
| OY | 1067 | CCTTAGTATATCTTAATAAAAAAATAATGCTGTCAACAAAPAGCGAGAGTGAATTAAGGA | 1126 |
| Db | 5981 | CAAAAGTAGACCTTTTGATGTATTAATGTTTATTTTCTTTAATTGACGACTTCATAG | 6040 |
| OY | 1127 | AAATTAATGGCGCCAAAGACTCTAAGGTATATTGACAAATCATTCAGAACTTTAAAAA | 1186 |
| Db | 6041 | ACCTCAAGTTGTGAAAGCTCAAGAAAGACTCCAAACGGTATTATTAATTAACATTACAAA | 6100 |
| OY | 1187 | GAAA 1190 | |
| Db | 6101 | TACA 6104 | |

```

RESULT 15
US-09-949-016-16067/c
; Sequence 16067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16067
; LENGTH: 50217
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(50217)

```


Mon Feb 28 08:49:24 2005

OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16067

| | | | | |
|-----------------------|-----------------|--------------------|------------|---------------|
| Query Match | 8.1%; | Score 404.8; | DB 4; | Length 50217; |
| Best Local Similarity | 76.2%; | Pred. No. 3.8e-84; | | |
| Matches 550; | Conservative 0; | Mismatches 112; | Indels 60; | Gaps 3; |

[illegible]

Search completed: February 27, 2005, 05:14:25
Job time : 797.948 secs

| | | | |
|----|------|---|------|
| Db | 1 | TCAGGAATCTGTAGAGTTCAGTTAGATCCAGTAAACCTTACCTGAGGCTGGGCTGTGC | 60 |
| OY | 4589 | CTTGAAGCTTGAAGCTGGCTTGAAGAGTGCCACTTTATTCTCCAGGCCCTTGCCTTGC | 4648 |
| Db | 61 | CTTGAAGCTTGAAGCTGGGCTTGAAGAGTGCCACTTTATTCTCCAGGCCCTTGCCTTGC | 120 |
| OY | 4649 | CCCTCAGATATCAGACACCCACCCCTCAGCTGGGCTGGGCTCTTGAAGTCAAAACCCAC | 4708 |
| Db | 121 | CCCTCAGATATCAGACACCCACCCCTCAGCTGGGCTGGGCTCTTGAAGTCAAAACCCAC | 180 |
| OY | 4709 | CCCCAGCCCAAGCCCGGCTCTGAGCCCGGCCAACCCATTTCGGTTCACAGACATGT | 4768 |
| Db | 181 | CCCCAGCCCAAGCCCGGCTCTGAGCCCGGCCAACCCATTTCGGTTCACAGACATGT | 240 |
| OY | 4769 | TCTGGCCAAACATATATTTTCTCTGTGACCAACCCCTCTTAAACCGTGCCCTCTGGAGCAA | 4828 |
| Db | 241 | TCTGGCCAAACATATATTTTCTCTGTGACCAACCCCTCTTAAACCGTGCCCTCTGGAGCAA | 300 |
| OY | 4829 | CCAGGACCTGGAGAGCTGGGGGCCGGGGAAAGACGCCCGGTCGATGACACAGACAGCCGAT | 4888 |
| Db | 301 | CCAGGACCTGGAGAGCTGGGGGCCGGGGAAAGACGCCCGGTCGATGACACAGACAGCCGAT | 360 |
| OY | 4889 | CATCAATGAATCCGACTGCGATATGACACACCGACCGGTGGACAGGCCGCTGTTGCTAAG | 4948 |
| Db | 361 | CATCAATGAATCCGACTGCGATATGACACACCGACCGGTGGACAGGCCGCTGTTGCTAAG | 420 |
| OY | 4949 | GCCCAACCAAGCTTATCTGGGGGGCGGTGTGTGATCAACAGTGGCTGCTC | 5000 |
| Db | 421 | GCCCAACCAAGCTTATCTGGGGGGCGGTGTGTGATCAACAGTGGCTGCTC | 472 |

```

RESULT 2
US-10-087-192-1642
; Sequence 1642, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1642
; LENGTH: 41454
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1642

```

| Query Match | 8.8% | Score 40.0 | DB 13 | Length 41454 |
|-----------------------|-------|--|---|------------------------------------|
| Best Local Similarity | 78.9% | Pred. No. 2.8e-94 | | |
| Matches | 569 | Conservative | 0 | Mismatches 133; Indels 19; Gaps 3; |
| QY | 345 | TTCTTTT | TTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTGTGTCGCCACAGC | 404 |
| Db | 37225 | TCCTTTGGTCTTCTTTT | TTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTGTGTCGCCACAGC | 3728 |
| QY | 405 | TGAGATGAGTGGCGGGATCTCGGCTCACTGCAAGCTCGGCTCCGGGTTACAGCATT | | 464 |
| Db | 37285 | TGGATGAGTGGCGGGATCTCGGCTCACTGCAAGCTCGGCTCCGGGTTACAGCATT | | 3734 |
| QY | 465 | CTCTGCTCAAGCCTCCCAAGTAGTGGGATCAACAGGGCCCGGCACTAAGCCCGGGTTAA | | 524 |
| Db | 37345 | CTCTGCTCAAGCCTCCCAAGTAGTGGGATCAACAGGACCCCGGCACTAAGCCCGGGTTAA | | 3740 |
| QY | 525 | T-----TTTTTTGTAATTTTAAAGACGGGGTTTACCGTTTTTACCCGGAGTGGCTCG | | 580 |

| | | | |
|----|-------|---|-------|
| Db | 37405 | TTTTATTTTTTTAATTTTAAAGTAGAGACGGGGTTTCACTGTTGTAGCCAGATGGTTTTCC | 37466 |
| Qy | 581 | ATCTCTGACCTCGTGATATCCGCCGCTTGCGCTCCCAAAGTGTCTGGGATTACAGGCTGT | 640 |
| Db | 37465 | ATCTCTGACCTCGTGATCCGCTGCGCTGCCCTCCCAAAGTGTCTGGGATTACAGGCTGT | 37524 |
| Qy | 641 | AGCCACCGGCGCGGCCCATGATCATCTTCTTGAC-----TAGCTGATGTGACAACT | 692 |
| Db | 37525 | AGCCACCGGCGCGGCCCTTTGTGCTTAACTTCTTTCGGCTTCATTTATTCAC | 37584 |
| Qy | 693 | ACCTAAGCATCAGACTTACCTTAA-----ATATGCAGTTTGGGCGAGCACCG | 745 |
| Db | 37585 | ATATTAATGACTGTGTCTCTTACCAAAATGCAGTATTAAAGAAATGAGGCGGCGCTCGG | 37644 |
| Qy | 746 | TGCTCATGCTGTAAATTCCAGACATTTTGGAGGCAAGGTGGTGTAATCACTTAAGGCC | 805 |
| Db | 37645 | TGCTCATGCTGTAAATCCAGACATTTTGGAGGCGAGGCGAGGCGAATCACTTAAGGTT | 37704 |
| Qy | 806 | AGAGTTTGAAGACGAGCTTGGCCAAATGTGAAACTCTGTCTTTAATAAAAAAAAAA | 865 |
| Db | 37705 | AGAGTTTGAAGACGAGCTTGGCCAAATGTGAAACCTTACTTTAATAAAAAAAAAA | 37764 |
| Qy | 866 | AAAAAAAAAAAAATACGCGGCGGTGTGCTGGGGCAACCTGTATATCCAGCTATGCTGAG | 925 |
| Db | 37765 | AAAAAACAATAAATTTAGCCAGGCAAGGTGCATGTCTGTATCTCCAGCCACTCGGGAG | 37824 |
| Qy | 926 | GCTGAGGCAAGAGTCACTTGAACCTGTGAGGCGGAGGTTGACGTGGGCGAGATCA | 985 |
| Db | 37825 | GCTGAGGCAAGAGATCACTTGAACCGAGAGGCGGAGGTTGACGTGAACCGAGATTGCA | 37884 |
| Qy | 986 | TCACCGGCTTTCAGGCTTGGGCGACAGACAAAGACTTGTCTCAATTAATTAATTAACA | 1045 |
| Db | 37885 | CCATTGCACTTCAGGCTGTGATGACAGAGTGAGACTCATCTCAAAAAAAAAAAAAAAAAA | 37944 |
| Qy | 1046 | A 1046 | |
| Db | 37945 | A 37945 | |

```

RESULT 3
US-10-300-611-4
/ Sequence 4, Application US/10300611
/ Publication No. US20040097451A1
/ GENERAL INFORMATION:
/ APPLICANT: Ming-Yi Chiang
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: MODULATION OF NIDOGEN EXPRESSION
/ FILE REFERENCE: PTS-0059
/ CURRENT APPLICATION NUMBER: US/10/300,611
/ CURRENT FILING DATE: 2002-11-19
/ NUMBER OF SEQ ID NOS: 136
/ SEQ ID NO 4
/ LENGTH: 91352
/ TYPE: DNA
/ ORGANISM: H. sapiens
US-10-300-611-4

```

| | | | | | |
|----|-----------------------|---|---------------------|-------------|----------------|
| | Query Match | 8.4% ; | Score 419 ; | DB 17 ; | Length 91352 ; |
| | Best Local Similarity | 79.4% ; | Prod. No. 4.5e-89 ; | | |
| | Matches 552 ; | Conservative 0 ; | Mismatches 125 ; | Indels 18 ; | Gaps 4 ; |
| QY | 350 | TTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGCTCTGTGCGCCAGCGTGAG | 409 | | |
| DB | 11334 | TTATTATTATTATTATTATTATTGAGATGAGCTTGCTGTGCGCCAGGTTGAG | 11393 | | |
| QY | 410 | TGACGTGGCGAGATCTGCGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCAATTCTCT | 469 | | |
| DB | 11394 | TGCAGTGGTGGCTCTGCGCTTACTCAAGCTCCACTCCGGGTTCCGCAATTCCT | 11453 | | |
| QY | 470 | GCCCTACACCTCCCAAGTAGTGGGAATTACAGGGCCCGCACTAAGCCCGGTAATTTT | 529 | | |
| DB | 11454 | GCCCTACACCTCCCAAGTAGTGGGAATTACAGCAACCCCGCACAGCCCGGCTAA-TTTT | 11512 | | |

475 AGCCTCCCAAGTAGCTGGGACTACAGGCGCCGCCACTACGCCCGGCTAATTTTTTGTGA 534

```

; SEQ ID NO 1450
;
; LENGTH: 26371
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens

```


[illegible]

```

Publication No. US20040005575A1
GENERAL INFORMATION:
APPLICANT: Roese et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C2
CURRENT APPLICATION NUMBER: US/10/227,577
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/091,504
PRIOR FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 09/764,869
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
Remaining prior art Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1945
LENGTH: 17397
TYPE: DNA
ORGANISM: Homo sapiens
US-10-227-577-1945

Query Match      8.3%; Score 415.2; DB 17; Length 17397;
Matches Local Similarity 78.6%; Pred. No. 2e+88;
Matches 558; Conservative 0; Mismatches 123; Indels 29; Gaps 4;

QY   TTTTGTATTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGACTCGCTGTCTGCCCGAGCTGGA    408
DB   15150 TGTTTGGAACTCTTTTGTGTGTGTATTAGAACAAGACTTGTCTGTGTGCCCAAGCTGGA    15209

QY   409 GTGCAGTGGCGGAGATTCGGCTCACATGCAAGCTCCGCTCCGGGTTCAAGCCATTCTCC    468
DB   15210 GTGCAGTGGCGACGATCTCGGCTCACATGCAAGCTCCGCTCCGGGTTCAAGCCATTCTCC    15269

QY   469 TGCTCTAGCTCCCAAGTAAGCTGGGACATAAGCGCCGCGCACATAAGCCGGCTAATTTT    528
DB   15270 TGCTCTAGCTCTCCGAGTAGCTGGGACATAAGCACCCGCGCACAAAGCCCGCTAA-TTT    15328

QY   529 TTGTGATTTTATTAAGACGCGGGTTTACCCTTTTATGCGGGATAGCATCGATCTCGT    588
DB   15329 TTGTGATTTTATTAAGACGCGGGTTTACCCTTTTATGCGGGATAGCATCGATCTCGT    15388

QY   589 ACCCTGTGATCCGCCCCGCTCGGCTCCCAAAGTGTGGATTTACAAGCGTAGCCACCG    648
DB   15389 ACCCTGTGATCTCGCCCCGCTCGGCTCCCAAAGTGTGGATTTACAAGCATGAGCTACTG    15448

QY   649 CGCCCGGCATGATCATCTTCTTGAATATGCTGATGTGACAGATACATAAGCCATCAGA    708
DB   15449 CACCCCGGC-----GATGTTTGCACCTCTTTAAACAGTGAACAAATCCCACCTTTTA--    15501

QY   709 CTGTACCTTTAATATATGACGTTTGGGACGAGGACGATGAGCTCATGCTGTAAATTCACG    768
DB   15502 -----TAAATATATTAAGCTTGGGCTGGGGCGCGTGGCTCACACCTGTAAATCCAGC    15553

QY   769 ACTTTGGAGGACAGAGTGCTGTGAATCATCTTAGGCGCAGAGATTGGAGACAAGCTTGCC    828
DB   15553 AGTTTGGAGGACCAAGGACAGTGATCATCTTAGGCTCGGCTGTTCAGAACCAAGCTTGCC    15612

QY   829 AACATGCTGAACCTGTCTTTTACTTAAAAAATAAAAAAAAAAAAAAAAAATACGCCGGGT    888

```



```

Db      15613 AACATGTTAAACGCTGCTTTACT-----AAAAATACAAAATTAGCCAGGT 15666
Oy      889 GTGCTGGGGCACACCTGTATATCCAGCTTANGCTGGAGGCTGAGGCACGAGAGTCACTTGA 948
Db      15661 GTGCTGGCACACACTGTAAATCCAGCTTACTGGGAGGCTGAGACACGAGAAATCACTTGA 15720
Oy      949 ACCCTGGAGGGGGAGGTTCAGGTGGGGGCGAGATCACTACACGCCCTCCAGCCTGGGGCA 1008
Db      15721 ACCGGGAGGGGAGAGGTTCAGTGAAGCGAGATTATGCCACTGCATCCACTTGGGGTA 15780
Oy      1009 CAGGCAAGACTGTCTCAATTAATTAATTAACAACAAGACAGCT 1058
Db      15781 AAGAGCAAGACTGTCTCAAAACATTAAAAATTAATTAATTAAGT 15830

RESULT 11
US-09-764-869-1943
; Sequence 1943, Application US/09764869
; Patent No. US20020061521A1
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1943
; LENGTH: 1933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-9-764-869-1943

```

```

Db      15606 AACATGTTAAACGCTGCTTTACT-----AAAAATCAAAAATTAAGCAAGT 156533
Qy      889  GTCGTGGGGCAACCTGTATATCCAGTACTGTGAGGCTGAGGCAAGAGTCACTTGA 948
Db      15654 GTGTGTGCACACACCTGTATATCCAGCTACTTGGGAGGCTGAAGCACGGAATCACTTGA 157133
Qy      949  ACCGTGAGGCGGAGGTTGCAGTGGGCGAGATCATATCACCGCCCTCCAGCTGGGCGA 1008
Db      15714 ACCCGGAGGCAAGAGTTGCAGTGAAGCCGAGATTATGCCATCTGCACTTCCACTTGGGTGA 157733
Qy      1009 CAGAGCAAGACTCTGTCTCAATTAATTAATTAACAAGCAACAGCACT 1058
Db      15774 AAGAGCAAGACTCTGTCTCAAAAATTAATAAATAATTAATTAAGT 15823

RESULT 12
US-10-091-504-1943
; Sequence 1943, Application US/10091504
; Publication No. US2003005908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ. ID NOS: 2442
; Prior Application Removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1943
; LENGTH: 19334
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1943

```

| Query Match | 8.3% | Score 415.2 | DB 9 | Length 19334 |
|-----------------------|----------------|---|-----------|--------------|
| Best Local Similarity | 78.6% | Pred. No. 2.1e-88 | | |
| Matches 558 | Conservative 0 | Mismatches 123 | Indels 29 | Gaps 4 |
| Oy | 349 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGCTCGCTCTGCGCCAGGCTGGA | 408 | |
| Db | 15143 | TGTTTCGTACCTCTTTTGTGTTGTTGTTAAAGACACACTTTGCTCTGTGCCACGGCTGGA | 15202 | |
| Oy | 409 | GTGACGTGCGGGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTACAGCCATTCTCC | 468 | |
| Db | 15203 | GTGACGTGCGACGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTACAGCCATTCTCC | 15266 | |
| Oy | 469 | TGCGTCAGACCTCCCAAGTACGTGGACCTAACAGCGCGCCGACCTACCGCGGTAAATTTT | 528 | |
| Db | 15263 | TGCGTCAGACCTCCCAAGTACGTGGACCTAACAGCGCGCCGACCTACCGCGGTAA-TTT | 15321 | |
| Oy | 529 | TTTGTATTTTAAAGTAGAGACGGGGTTACCGTTTATGCGGGATGGCTCTGATCTTCG | 588 | |
| Db | 15322 | TTTGTATTTTAAAGTAGAGACGGGGTTTGGCGTTTATGACAGATGGTCTGATCTTCG | 15381 | |
| Oy | 589 | ACCGTGTATCCGCGCGCTCGGCTCCCAAAAGCTGGGATTTACAGCGGTGAGCCACCG | 648 | |
| Db | 15382 | ACCTGTGATTTGCCCGCTCGGCTCCCAAAAGCTGGGATTTACAGGCAATGAGCTACAG | 15441 | |
| Oy | 649 | CGCCCGGACGATCATCTTCTTCACTATCGTAGTGAACAGTACTTAAGCATCAGA | 708 | |
| Db | 15442 | CACCCGGC-----GATGTTTGCACTCTTTAACAGTGAACAAATCCAGCTTTT-- | 15499 | |
| Oy | 709 | CTCTAACCTTAAATATGACATTTGGGACAGGACCGTGGCTCATGCTGTAAATTCACG | 768 | |
| Db | 15495 | -----TAAATATATAAGCTGGGCGCGGTGGCTCACACTGTATCCACG | 15545 | |
| Oy | 769 | ACTTTGGGAGGACAGGTGGGTGAATCACTTAGGACGAGATTAGAACCAAGCTGGCC | 828 | |
| Db | 15546 | AGTTTGGGAGGACCAAGGACAGGTGATCACTTAGAGCTGGGTGTTCAAGACCACTTGGCC | 15605 | |
| Oy | 829 | AACATGTGGAACTCTGCTTTACTTAAAAAATTTTAAAAAATTTTAAAAAATTCAGCCGGGT | 888 | |

| | Best Local Similarity | 78.64; | Pred: No. 2,1e-88; | |
|----|-----------------------|--|--------------------|---------------------------------------|
| | Matches | 558; | Conservative | 0; Mismatches 123; Indels 29; Gaps 4; |
| QY | 349 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGAGACTCGCTCTGTCCGCCAGGCTGGA | 408 | |
| Db | 15143 | TGTTTCGTAAGCTCTTTTTTGTGTGTGTTTAAAGACAGACTCTTGCTCTGTCCGCCAGGCTGGA | 15202 | |
| QY | 409 | GTCGAGTGGCGGGAATCTCGGCTCAGTCGAAAGCTCCGCTCCCGGGTTCAAGCAATCTTCC | 468 | |
| Db | 15203 | GTGCAAGTGGACAGATCTCGGCTCAGTCGAAAGCTCCGCTCCCGGGTTCAAGCAATCTTCC | 15262 | |
| QY | 469 | TGCTCTCAGCTCCCAAGTAGCTGGAGCTACAGGCGCCGCTCAGTACGCCCCGTAAATTT | 528 | |
| Db | 15263 | TGCTCTCAGCTCCCGAGTAGCTGGAGCTACAGGCAACCCGCTCAGAGCCCGGTAA-TTT | 15321 | |
| QY | 529 | TTTTGTAATTTTAAAGAGACGGGGTTTCAACCTTTTAAAGCGGAGTGGGCTCGATCTCCG | 588 | |
| Db | 15322 | TTTTGTAATTTTAAAGAGAGGGGGTTTGGCGGTAAAGCAGAGATGCTCGAATCTCCG | 15381 | |
| QY | 589 | ACCTCGTAGTCGCGCCGCTCGGCTCCCAAAAGTGTGGGAATTAAGCGGTAGCCACCG | 648 | |
| Db | 15382 | ACCTGTGATCTGCGCCCTCGGCTCCCAAGTGTGGGAATTAAGCGGTAGCTACG | 15441 | |
| QY | 649 | CGCCCGGCAATGATCAATCTTCTTGAAGTAGCTGATGTAAGATCACTTAAAGCAATCAGA | 708 | |
| Db | 15442 | CACCCGGC-----GATGTTTGCAATCTTTTAAAGTAAAGAAATCCAGCTTTTA-- | 15494 | |
| QY | 709 | CTCTACCTTTAAATATGACGTTTGGGACAGGACCGTGGCTCAGCTGTAAATCCAGC | 768 | |
| Db | 15495 | -----TAAATATATATAGCTGGGCTGGGGCGGTTGGCTCAACCTGTAAATCCAGC | 15545 | |
| QY | 769 | ACTTTGGGAGGACAGGTGGGTGAATCACTTGAAGGCAAGAGTTTGAACAGGCTGGCC | 828 | |
| Db | 15546 | AGTTTGGGAGGACAGGCAAGGCAAGTGAATCACTTGAAGCTCGGGTGTTCAGAGCCAGCTTGGCC | 15605 | |
| QY | 829 | AACATGTGAAGACCTGCTTTTAACTAAAGAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCGGGT | 888 | |
| Db | 15606 | AACATGTGAAGACCTGCTTTTAACT-----AAAAATACAAAATTAAGCAAGCT | 15653 | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 20:26:59; Search time 15671.4 Seconds
(without alignments)
12144.524 Million cell updates/sec

Title: US-09-936-271B-13_COPY_1_5000

Perfect score: 5000
Sequence: 1 gggcccccagctgaagcgaag.....tgcattccacagtgcctc 5000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|----------------------|
| C 1 | 397.6 | 8.0 | 2971 | 8 AF101960 | AF101960 AF101960 |
| C 2 | 373.8 | 7.5 | 2971 | 8 AF101960 | AF101960 AF101960 |
| C 3 | 365 | 7.3 | 1715 | 3 CRS92231 | CRS92231 full-length |
| C 4 | 362.6 | 7.3 | 2429 | 3 HSM805366 | AL834319 Homo sapi |
| C 5 | 337.4 | 6.7 | 652 | 1 AU120416 | AU120416 Homo sapi |
| C 6 | 337.2 | 6.7 | 1641 | 3 BC033224 | BC033224 Homo sapi |
| C 7 | 336 | 6.7 | 2772 | 3 BC038630 | BC038630 Homo sapi |
| C 8 | 334.4 | 6.7 | 698 | 5 BM999040 | BM999040 UI-H-D10 |
| C 9 | 331.8 | 6.6 | 5528 | 9 AB012089 | AB012089 Homo sapi |
| C 10 | 329 | 6.6 | 381 | 8 CC061638 | CC061638 MUGQ CH25 |
| C 11 | 325.2 | 6.5 | 652 | 1 AU120416 | AU120416 Homo sapi |
| C 12 | 324 | 6.5 | 701 | 2 BR744242 | BR744242 601576428 |
| C 13 | 324 | 6.5 | 5511 | 3 HSM807323 | BK447179 Homo sapi |
| C 14 | 322.8 | 6.5 | 6934 | 3 HSM807092 | BK440943 Homo sapi |
| C 15 | 321.6 | 6.4 | 2772 | 3 BC038630 | BC038630 Homo sapi |
| C 16 | 318.6 | 6.4 | 5528 | 9 AB012089 | AB012089 Homo sapi |
| C 17 | 315.6 | 6.3 | 709 | 9 AG010128 | AG010128 Homo sapi |
| C 18 | 315 | 6.3 | 710 | 6 CA428305 | CA428305 UI-H-DP0 |
| C 19 | 312.5 | 6.3 | 542 | 1 A1679782 | A1679782 tU76E10.x |
| C 20 | 312.2 | 6.2 | 495 | 1 A1963720 | A1963720 w76B04.x |
| C 21 | 311.8 | 6.2 | 2429 | 3 HSM805366 | AL834319 Homo sapi |
| C 22 | 309.6 | 6.2 | 3491 | 3 BC039100 | BC039100 Homo sapi |
| C 23 | 309.6 | 6.2 | 1910 | 3 BC035771 | BC035771 Homo sapi |
| C 24 | 309.4 | 6.2 | 3491 | 3 BC039100 | BC039100 Homo sapi |

| | | | | | |
|------|-------|-----|------|-------------|--------------------|
| C 25 | 309.2 | 6.2 | 861 | 1 AL558859 | AL558859 |
| C 26 | 308.4 | 6.2 | 764 | 3 BU617735 | BU617735 UI-H-DP0 |
| C 27 | 305.4 | 6.1 | 2076 | 3 BC035510 | BC035510 Homo sapi |
| C 28 | 305 | 6.1 | 834 | 8 B2612098 | B2612098 WHADP72TR |
| C 29 | 304.2 | 6.1 | 6934 | 3 HSM807092 | BK440943 Homo sapi |
| C 30 | 303.6 | 6.1 | 460 | 2 BF033594 | BF033594 601453571 |
| C 31 | 302.8 | 6.1 | 322 | 5 BU941317 | BU941317 AGENCOURT |
| C 32 | 301.6 | 6.0 | 328 | 1 AL696113 | AL696113 DKFZP686A |
| C 33 | 301.2 | 6.0 | 868 | 2 BP337291 | BP337291 602034658 |
| C 34 | 300.2 | 6.0 | 658 | 3 CN273613 | CN273613 17006000 |
| C 35 | 299.8 | 6.0 | 1073 | 8 B2601227 | B2601227 WHADP89TR |
| C 36 | 299.8 | 6.0 | 1910 | 3 BC035771 | BC035771 Homo sapi |
| C 37 | 299.8 | 6.0 | 2076 | 3 BC035510 | BC035510 Homo sapi |
| C 38 | 299.6 | 6.0 | 829 | 5 B2609393 | B2609393 WHAC148TF |
| C 39 | 299.4 | 6.0 | 599 | 5 BU729004 | BU729004 UI-E-C01 |
| C 40 | 299.2 | 6.0 | 691 | 1 AL043009 | AL043009 DKFZP4340 |
| C 41 | 299 | 6.0 | 1366 | 3 HSM800180 | AL050097 Homo sapi |
| C 42 | 298.8 | 6.0 | 584 | 4 BM701794 | BM701794 UI-E-C01 |
| C 43 | 298.6 | 6.0 | 698 | 5 BM999040 | BM999040 UI-H-D10 |
| C 44 | 297.6 | 6.0 | 844 | 8 A0748723 | A0748723 HS 5540 A |
| C 45 | 297.2 | 5.9 | 5598 | 3 HSM803450 | AL832143 Homo sapi |

ALIGNMENTS

| | | | | | | | |
|---|------------|--|--|---------|-----|--------|-----------------|
| RESULT 1 | AF101960/c | AF101960 | Human Homo sapiens genomic clone PTWB59.14, genomic survey sequence. | 2971 bp | DNA | linear | GSS 06-NOV-2000 |
| LOCUS | AF101960 | Human Homo sapiens genomic clone PTWB59.14, genomic survey sequence. | | | | | |
| DEFINITION | AF101960 | Human Homo sapiens genomic clone PTWB59.14, genomic survey sequence. | | | | | |
| ACCESSION | AF101960 | AF101960.1 | GI:4193786 | | | | |
| VERSION | AF101960 | AF101960.1 | GI:4193786 | | | | |
| KEYWORDS | AF101960 | AF101960.1 | GI:4193786 | | | | |
| SOURCE | AF101960 | AF101960.1 | GI:4193786 | | | | |
| ORGANISM | AF101960 | AF101960.1 | GI:4193786 | | | | |
| REFERENCE | AF101960 | AF101960.1 | GI:4193786 | | | | |
| AUTHORS | AF101960 | AF101960.1 | GI:4193786 | | | | |
| TITLE | AF101960 | AF101960.1 | GI:4193786 | | | | |
| JOURNAL | AF101960 | AF101960.1 | GI:4193786 | | | | |
| MEDLINE | AF101960 | AF101960.1 | GI:4193786 | | | | |
| PUBMED | AF101960 | AF101960.1 | GI:4193786 | | | | |
| COMMENT | AF101960 | AF101960.1 | GI:4193786 | | | | |
| FEATURES | AF101960 | AF101960.1 | GI:4193786 | | | | |
| source | AF101960 | AF101960.1 | GI:4193786 | | | | |
| ORIGIN | AF101960 | AF101960.1 | GI:4193786 | | | | |
| Query Match | AF101960 | AF101960.1 | GI:4193786 | | | | |
| Best Local Similarity | AF101960 | AF101960.1 | GI:4193786 | | | | |
| Matches 552; Conservative 0; Mismatches 119; Indels 43; Gaps 4; | AF101960 | AF101960.1 | GI:4193786 | | | | |
| 342 ATCTCTTTT | AF101960 | AF101960.1 | GI:4193786 | | | | |
| 1141 ATCCAGTTTC | AF101960 | AF101960.1 | GI:4193786 | | | | |
| 402 GGCTGAGTC | AF101960 | AF101960.1 | GI:4193786 | | | | |

| | | | | |
|-----------------|--|--|--|------------------------|
| Db | | 1081 | G6CTGGAATGCAGTGGCANTGATCTTTGGGTCACTCGAAGCTCCGCTCCTGGATTCAACACC | 1022 |
| OY | | 462 | ATTCTCTGTCCTCAAGCTCTCCCAAGTAGCTGGAATCAAGCGGCCCGCATCAAGCCCCGAC | 521 |
| Db | | 1021 | ATTCTCCCGCTCAAGCTCTCCGAGAAGCTGGGAATCAAGATGCCACAACATCGAACCCAGC | 962 |
| OY | | 522 | TAAATTTT---TTTGATATTTTAGTAGAACAAGGGGTTTCAACGTTTTAGCCGGAGTGGCT | 578 |
| Db | | 961 | TAAATTTTAAATTTTAATTTTAATTTTAGTAGAACAAGGGGTTTCAACGTTTTAGCCGGAGTGGCT | 902 |
| OY | | 579 | CGATCTCCTGACCTTCGTGATCCGCGCCGCTCGGCTCCCAAAGTCTGGGATTCAGAGCG | 638 |
| Db | | 901 | CGATCTCCTGACCTTCGTGATCCGCGCGCTCGGCTCCCAAAGTCCAGAGATTCAGAGCG | 842 |
| OY | | 639 | TGAGCCACCGCGCCCGCCGATGATCACTTCTTGAATACTGATGTGACAAGTAAGCTTAA | 698 |
| Db | | 841 | TGAGCCACTGTGCTGGC-----TGTTCTGATCTTAGA | 808 |
| OY | | 699 | AGCATCAAGACTCACCTTTAAATATGACAGTTTGGCCGAGCAACCGGCTCATGCTCG | 758 |
| Db | | 807 | AACATTCAGTCTTTCACCATTAATTAATTAATGTTGGGCCAATATGGGTGCTTAATGCTG | 748 |
| OY | | 759 | TAAATTCAGCACTTTGGAGGCGAGAGGTGGTGAATCACTTGAGGCCAGAGTTTGAAGAC | 818 |
| Db | | 747 | TAAATCCAGCACTTTGGAGGCTGAGGTGGGTGATCAAC--GAGTCAAGAGTTTGAAGAC | 690 |
| OY | | 819 | CAGCTGGCCAAATGATGTAAACTCTGTCTTAACTAAAAAAAAAAAAAAAAAAAAAAA | 878 |
| Db | | 689 | CAGCTGGCCAAATGATGTAAACTCATCTCACT-----AAATATCAAAAA | 642 |
| OY | | 879 | TCATCCGGGTGTCTGTGGGGCACACTGTAAATCCCAAGCTATGCTGAGGCTTAGGCACGAG | 938 |
| Db | | 641 | TTAGTGTGTGTGTGTGGGCGGCGCTGTAAATCCAGCTCACTTGGAGGCTTAGGCACGAG | 582 |
| OY | | 939 | AGTCACTTGAACCCCTGGAGGCGAGAGTTGAGCTGAGGCGCAGATCAATCAACGCTCTCCA | 998 |
| Db | | 581 | AATCGCTTGAACCCAGAGGCGGAGAGTTGACAGTGAAGCCAAATGTGTGCATCGACATCCA | 522 |
| OY | | 999 | GCTTGGGCGACAGACAAAGACTCTGTCTCAATAATAATAATPAAACAAACGACAA | 1052 |
| Db | | 521 | GCTTGGGCGACAGAGTAGACATCTGTCTCGAATAATAATAATAATAATAATAATAA | 468 |
| <hr/> | | | | |
| RESULT 2 | | | | |
| AFI01960 | | 2971 bp | DNA | linear GSS 06-NOV-2000 |
| LOCUS | | AFI01960 Human Homo sapiens genomic clone PTM59.14, | | genomic survey |
| DEFINITION | | sequence. | | |
| ACCESSION | | AFI01960 | | |
| VERSION | | AFI01960.1 | | GI:4193786 |
| KEYWORDS | | GSS. | | |
| SOURCE | | Homo sapiens (human) | | |
| ORGANISM | | Homo sapiens | | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| REFERENCE | | 1 (bases 1 to 2971) | | |
| AUTHORS | | Begler,G., O'Brian,K.C., Kim,Y.C., Schreiber,G. and Pitterle,D.M. | | |
| TITLE | | A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH1A metastasis suppressor region | | |
| JOURNAL MEDLINE | | Genomics 55 (2), 164-175 (1999) | | |
| PUBMED | | 99134294 | | |
| COMMENT | | Contact: Begler G Medicine and Radiology Duke University Medical Center Box 2610, MSRB, Room 117, Durham, NC 27710, USA part of a 1.4 megabase contig including the LOH1A metastasis suppressor region Bin T Class: unknown. Location/Qualifiers 1..2971 /Organism="Homo sapiens" | | |

```

/mo1_type="genomic DNA"
/mab_xref="taxon:9606"
/ab_xref="1p15.5"
/clone="PTWB59.14"
/clone_1fb="Human"

```

| | | | | |
|-----------------------|--------------|--------------------|-----------------|--------------|
| Query Match | 7.5%; | Score 373.8; | DB 8; | Length 2971; |
| Best Local Similarity | -71.2%; | Pred. No. 6.7e-23; | | |
| Matches 566; | Conservative | 0; | Mismatches 192; | Indels 37; |
| | | | Gaps | 4. |

[illegible]

ORIGIN

CLRDWGCPRRAARAAGAL "

| | | | | |
|---------------------------|-------|--------------------|------------|--------------|
| Query Match | 7.3% | Score 362.6; | DB 3; | Length 2429; |
| Best Local Similarity | 75.3% | Pred. No. 6.4e-22; | | |
| Matches 531; Conservative | 0; | Mismatches 114; | Indels 60; | Gaps 4 |

[illegible]

| | | | | | | |
|------------|------------|--|--|------|--------|-----------------|
| RESULT 5 | AU120416/c | AU120416 | 552 bp | mRNA | linear | EST 01-AUG-2002 |
| LOCUS | | AU120416 | | | | |
| DEFINITION | | AU120416 | HEMBB1 Homo sapiens CDNA clone HEMBB1000637 5', mRNA sequence. | | | |
| ACCESSION | | AU120416 | | | | |
| VERSION | | AU120416.1 | GI:10935651 | | | |
| KEYWORDS | | EST. | | | | |
| SOURCE | | Homo sapiens (human) | | | | |
| ORGANISM | | Homo sapiens | | | | |
| REFERENCE | | EukaryaOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 652) | | | | |
| AUTHORS | | Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Iisaga, T. | | | | |
| TITLE | | HRI human cDNA project | | | | |

JOURNAL
COMMENT
Unpublished (2000)
Contact: Takao Isogai

FEATURES

Location/Qualifiers
1. .652

ORIGIN

| | | | | |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match | 6.7% | Score 337.4; | DB 1; | Length 652; |
| Best Local Similarity | 75.7% | Pred. No. 1.8e-19; | | |
| Matches 494; Conservative | 0; | Mismatches 125; | Indels 34; | Gaps 5 |

| | | | |
|----|-----|---|------|
| QY | 400 | AAGCCTGAGATGAGAGTGGCGGGAACTCGGGCTCACTGCAAGGTCGCGCTCCGGGTTTCAAG | 459 |
| Db | 642 | CAGNTTGGAGTGCATAAGCGATGATCTCAAGCTCACTGCAACCTCTGCTCCCGGTTTCAAG | 583 |
| QY | 460 | CCATTCTCTGGCTCTCAGCGCTCCCAAGTAGCTAGGAGTACAGAGCGCCGCACTACGCCCG | 519 |
| Db | 582 | CGAATTCCTCGGCTCAAGCTCTCTGAGTAGCTGGAGCTTCAGATGAGCAACCATGCCCA | 523 |
| QY | 520 | GCTA-----ATTGTTTGTATTTTGTAGTAGAGCGGGGTTTCACCGTTTACGCCGG | 571 |
| Db | 522 | GCTATCTTTTNTTTTTTTTTTTTTTTTTTAAAGTAGAGACAGGGGTTTTCACGTGTATGACCAAG | 463 |
| QY | 572 | ATGGCCTCGATCTCTCTGACCTCGTGAATCCGCGCCGCTCGGCGCTCCCAAGTCTGGAGTT | 631 |
| Db | 462 | ATGGTCTCGATCTCTGATCTCGTAGTCTCGTAGTGGCCACTCGGTTCCCAAGTCTGGAGTT | 403 |
| QY | 632 | ACAGCGGTGAGCCACCGCGCCCGGCC--ATGATCATCTTCTTGACTATGCTGATGTGA | 689 |
| Db | 402 | ACAGCGCTGAGCCACAGACCCGCGCCGATATATATATTTGTGATTTTAAAAAGGCC | 343 |
| QY | 690 | AGTACCTTAAAGCCATCAGACTTACCTTTAAATATGATGATTTGGGCGCAGGCAACCGTGC | 749 |
| Db | 342 | AAGTCAGAAATCAAAACAACAACAATAA-----GCGCAGGACCGTGGC | 294 |
| QY | 750 | TCATGCTCTGTATTTCCAGCACTTTGGAGGAGGAGGTGGTGAATCACTTGAAGCCAGGA | 809 |
| Db | 293 | TCAGGCTCTGTATCTCCAGCACTTGGAAAGGCCAGGCGAGTTGATTCACCTGAGGTCAAGA | 234 |
| QY | 810 | GTTTGAACAACAGCTCGGCGCAACATGTGTGAATCTGTCTTTATCTTAAAAA | 869 |
| Db | 233 | GTTTGAACAACAGCTCTTCCAAACATGTGTGAACCCCATCTTCACT-----AAAA | 186 |
| QY | 870 | AAAAAAAATCAGCGCGGAGTGTCTGTGGGGCAACCTGTATATCCAGCTATGTGTGAGAGCTG | 929 |
| Db | 185 | ATACAAAAAATTAGCTGGGCGGTGATGTCAAGTGCTCTGTATCTCCAGCTACTCTGAGGCTG | 126 |
| QY | 930 | AGGACAGAGATCATTTGAACCTTGAAGCGCGAGGTTTGCAGTGGGCGAGATCACATCAC | 989 |
| Db | 125 | AGGTAGAGAGATTGCTTGAACCCAGGAGGTGGAAGTTGCAACGAGCGCAGAGATCGGCCAT | 66 |
| QY | 990 | CGCCCTCAGCCTGGGGGAC--AGAGCAAGACTCTGTCTCAATAATAAATTA | 1041 |
| Db | 65 | TGCATCTCAGCTGGAACGGCAAGGCAAAATCTGTCTCAAAAAA | 13 |

RESULT 6
BC033224/c
LOCUS BC033224 1641 bp mRNA linear HTC 27-JUN-2002
DEFINITION Homo sapiens, clone IMAGE:5022014, mRNA.
ACCESSION BC033224
VERSION BC033224.1 GI:21619936
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1641)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
infodbcgc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susandha Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Cloned distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 43 Row: 1 Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
This clone has the following problem: frame shifted.

FEATURES
source location/Qualifiers

1..1641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5022014"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 6.7%; Score 337.2; DB 3; Length 1641;
Best Local Similarity 75.3%; Pred. No. 1.1e-19;
Matches 500; Conservative 0; Mismatches 128; Indels 36; Gaps 5;
QY TTTTGTATTTTTCAGTAGAGACAGGGTTTCACCTGTTAGCCAGGATGCTCGATCTCT 407
DB TTTTGTATTTTTCAGTAGAGACAGGGTTTCACCTGTTAGCCAGGATGCTCGATCTCT 1059
QY 408 AGTCAGATGGGGGATCTCGGCTACAGCAAGCTCCCGGGTTTCACGCAATCTTC 467
DB 1058 AGTCAGATGGGGGATCTCGGCTACAGCAAGCTCCCGGGTTTCACGCAATCTTC 999
QY CTGCTCAGGCTCCCAAGTAGCTGGGACTACAGGCGCCCGCACTAGCCCGGCTAATT 527
DB 998 CTGCTCAGGCTCCCAAGTAGCTGGGACTACAGGCGCCCGCACTAGCCCGGCTAATT 940
QY TTTGTATTTTTCAGTAGAGACAGGGTTTCACCTGTTAGCCAGGATGCTCGATCTCT 587
DB TTTGTATTTTTCAGTAGAGACAGGGTTTCACCTGTTAGCCAGGATGCTCGATCTCT 587

DB TTTGTATTTTTCAGTAGAGACAGGGTTTCACCTGTTAGCCAGGATGCTCGATCTCT 880
QY 588 GACCTGCTATTCGCGCGCTCGGCTCCCAAGTGTGGGATTAACAGGGCTGAGCCACC 647
DB 879 GACCTGCTATTCGCGCGCTCGGCTCCCAAGTGTGGGATTAACAGGGCTGAGCCACC 820
QY 648 GCGCGCGCGCGATGATCTCTTCTTGAATGATGATGATGATGATGATGATGATGAT 707
DB 819 ATGCCAGGCTTAATCTCTATTTGAATGATGATGATGATGATGATGATGATGATGAT 761
QY 708 ACTCTACCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 767
DB 760 -----CTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 709
QY 768 CACTTTGGAGGCGAGAGTGGTGAATCACTTTGAGGCGCAGAGATTTGAGACCACTTGGC 827
DB 708 CACTTTGGAGGCGAGAGTGGTGAATCACTTTGAGGCGCAGAGATTTGAGACCACTTGGC 650
QY 828 CAACATGATGAATCTGCTTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 887
DB 649 CAACATGATGAATCTGCTTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 602
QY 888 TGTGTTGGGCGACACCTGTAATCCAGCTATGCTGAGGCTGAGGCGACGAGATCACTTG 947
DB 601 ACCTGGGCGACCGCGCTGTAATCCAGCTATGCTGAGGCTGAGGCGACGAGATCACTTG 555
QY 948 AACCTGGAGGCGAGGTTGCTGAGGCGAGATCACTATCACTATCACTATCACTATCACTAT 1007
DB 554 AACCTGGAGGCGAGGTTGCTGAGGCGAGATCACTATCACTATCACTATCACTATCACTAT 495
QY 1008 ACAG 1011
DB 494 ACAG 491

RESULT 7
BC038630/c
LOCUS BC038630 2772 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, Similar to hypothetical protein FLJ20489, clone IMAGE:5263792, mRNA.
ACCESSION BC038630
VERSION BC038630.1 GI:24116283
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2772)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadnan@systemsbiology.org
Anup Madan, Jessika Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Cloned distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 73 Row: 1 Column: 4
This clone has the following problem: retained intron.
location/Qualifiers
1..2772

| | | | |
|------------|---|---|--|
| Db | 241 | CTCTGACACTGTATCTGCCGCTCCACTCCCAAGTCTGGGATTACAGCGCTGAG | 300 |
| OY | 643 | CCACCGCGCCCGGCCATGATCATCTTCTTGACTGATGTGCAAGTACTTAAAGCC | 702 |
| Db | 301 | CCACCATCCCTGGCCATTATATTTCTTTTCTTTTCGTTTAATCAGCAATGACCA | 360 |
| OY | 703 | ATCAGACTCTACCTTTAAATATGAGTTTGGGCCAGGCACCGTCCTCATGCCCTGAT | 762 |
| Db | 361 | GAAGACTGACGAGACTACT-----GTTGGCCGGCGTGTGACTCATATCTGAAT | 414 |
| OY | 763 | TCCAGCATTTTGGAGCAGAGGTGGTGAATCACTTGAGGCCAGAGTTTGAAGCCAGC | 822 |
| Db | 415 | CCAGTATTTCAGGAAGCAAGAGGGTGAATTTTGTAGCCGGAATTTGAGACCAAG | 474 |
| OY | 823 | CTGGCCAACTGTGTGAACTCTGTCTTTACTTAAAAAATTTTTTTTTTTTTTATCAG | 882 |
| Db | 475 | CTGGGCACAATGGCAAAAACCTGTCTCTACAAAAATTA----- | 513 |
| OY | 883 | CCGGGTGTGTGGGGCACACCTGTAACTCCAGCTATGCTGAGGCTGAGGACAGAGTC | 942 |
| Db | 514 | ---GGTGTGTGTCCACACGCTGTGTGTCCAGCTACTGGAGGCTGAGTATGAGCTC | 570 |
| OY | 943 | ACTTGAACCTTGAGGCGGAGGTTCAGATGGGCGGAGATCATCAATCCGCCCTCAGCT | 1002 |
| Db | 571 | ACTGAACCTTGGAGGTAAAGGCTGCAATGAGCGAGATCTCGACACTGTACTCAGCT | 630 |
| OY | 1003 | GGGGGACAGACG-AGATCTGTCTCAATTAATTAATTAATTAACCAACGACAGCTTTG | 1061 |
| Db | 631 | GAGTGGCAGACCAACCTATCTCANAAATATTTTANNAAGAAAAAATTTTTAAAAA | 690 |
| OY | 1062 | TTGTACC 1068 | |
| Db | 691 | TCGTGCC 697 | |
| RESULT 9 | | | |
| LOCUS | AB012089 | 5528 bp | DNA linear GSS 21-MAY-1998 |
| DEFINITION | AB012089 | | Human sapiens gene, genomic survey sequence. |
| ACCESSION | AB012089 | | |
| VERSION | AB012089.1 | | GI:3149956 |
| KEYWORDS | GSS. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | 1 (state) Kurashashi,H., Sakamoto,M., Ono,J., Honda,A., Okada,S. and Nakamura,Y. | | |
| TITLE | Molecular cloning of the chromosomal breakpoint in the LIS1 gene of a patient with isolated lissencephaly and balanced t(8;17) | | |
| JOURNAL | Hum. Genet. (1998) in press | | |
| REFERENCE | 2 (bases 1 to 5528) Kurashashi,H. | | |
| AUTHORS | Direct Submission Submitted (09-MAR-1998) Hiroki Kurashashi, Biomedical Research Center, Osaka University Medical School, Division of Clinical Genetics, Department of Medical Genetics, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:kura@cigene.med.osaka-u.ac.jp, Tel.:81-6-879-3381, Fax:81-6-879-3389) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..5528 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17p13.3" | | |
| ORIGIN | | | |

RESULT 10
CC061638/c

CC061638 381 bp DNA linear GSS 28-Apr-2003

3299 AAGAACCAAGCGCGGGCGGGGCGGTGGCTACACCGCTGTAATCCAGCCCTTTTGGAGGCGCAGG 3358
|||
2123 AAAATATACAGGCGAGGCGGGGTGGCTCACGCGCTGTAATCCAGCACTTGGGGAGGCGCAGG 2182
|||
3359 CAGGCGGATCACTGAGGTCGGGAGATCGAGGCCAGGCTGACCAACATGAG-AAACCCG 3417
|||
2183 CAGGTGATTCACCTGAGGTCAGAGGTTTGAGAACAGCGCTGGCTTAACATGGTGAACCCCG 2242
|||
3418 TCTCTACTAAAAATATCAAAAAATTAGCTGTGTGTGGCGCGTGCCTGTAATCCAGCTA 3477
|||
2243 TCTCTACTAAAAATATCAAAAAATTAGCTGTGTGTGGCGGTGTGATCCTGTAAATCCAGCTA 2302
|||
3478 CTGGGAGACCTGGGGGACGAAAAATACGCTGTAACCCGGGAGGTGAGGTTGGGTGAGCG 3537
|||
2303 CTGGGAGGCTGAGGCGAGAGAAATTGCTTGAAACCGAGGACAGAACTTGACATGATTTG 2362
|||
3538 AGATCGCGCCATTGCACTCCAGCGCTGGGCAACAAGCGAACTCCGCTTCGAAGAAGAA 3597
|||
2363 AGATTGACACCACTGATCTCGAGCGTGGGCAAAAAAGAGAAAATTGTGTTCMAGGAAA 2422
|||
3598 AAAGAA-----AAAAAAAAAGGTAAACCACTGATGGGC 3633
|||
2423 AAAAAAGTATGTATATATATATATACACACACATTTTGTATTATCAATTTCAATAGA 2482
|||
3634 ACGGAGGACTGATGATGAGAGTGGGGGACATGATGTAGCTGTAGCTGTGTGAGAGG 3693
|||
2483 AGGGGAAAAATCTTGATTTGATTCATTGGGTAATTTAAATTTTCTCAGATCTTTTTTTT 2542
|||
3694 AGGAGATTGACAGGATTGAGAAAGGCATGTTTTCATCTGAGAAATTCAAGAACTTAGCCCTG 3753
|||
2543 TTTTTCGTGACACGGAAGTCTCTCTGTGCGCTTAGCTGAGGTGACAGTGGCAGCATTTTG 2602
|||
3754 CTCTTCCCTCATGTGTGCGCCCTTAAGCTGAGCCCTTCT-----TTCTGTGTCTGC 3805
|||
2603 GCTTATCTGCAACCTCCACCTCCAGGTTTCAGTGTCTTCTGCTCAGCCTCTCTGATAGCT 2662
|||
3806 TTTTGGAAACCTTAGCTCCGCGCCATGAGCTGACCCCACTCTTCTCAACACGCCC 3865
|||
2663 GGGATTACAGGCAATGGCGAACACGCGCGGTAAATTTTGTATTTTAATGTAAGATGGGG 2722
|||
3866 CTAGGCGACACTAGTGAAGCCCGCCTAAGGCCAACCCCTT-----TGGGCGACGCTC 3920
|||
2723 TTTCAACAATGTGGTCAAGACTGATGTCTCAAACTCTGACCCCGTATCTCCACTTGGC 2782
|||
3921 CACCCCTATCTGTGGGTACCTTTGAAACCCCTTCAAAAGTCAGAGCTTTTTTTTTT 3980
|||
2783 CTCCCAAGTGTGGGATTACAGGCGTGAAGCAATCGCGCTGGCAGATTATTTTATTT 2842
|||
3981 TTTTTTTGGAGACAGTCTTGTCTCTCTCCCAAGGCTGGAGTGCAGTGGCGTGTACTCGGC 4040
|||
2843 TCTTTTTTGAGACGGAATGTAAGTCT--TGTGCCAAGAGTGCATAAGCGGATCTTGTGC 2901
|||
4041 TCACCTCAACCTTGCCTCCAGGTTCAAGTATTTCTGTCTCCACTCTCTGATGAC 4100
|||
2902 TCACCTGCAACTCCACCTCCAGGTTCAAGAGATTTTCTGTCTCAACCTCTCCAGTATGC 2961
|||
4101 TGGGATTACAGGTGCGGGCACCAACGCTGGCTAATTTTGTGTCTTATGAGAGACAGG 4160
|||
2962 TGGGATTACAGGTGTGTGTCACCAACCCCGCTAATTTTGTATTTTATGAGAGATGGA 3021
|||
4161 GTTTCACCTTGTGGCAGGCTGTGTCTCAAACTCCCAACCTCAGGTATCCGCCACCTC 4220
|||
3022 GTTTCACCAATTTGGCAGGCTGTGTCTTCCAAACCTCAGCTCAAGTATCCCGCGCTC 3081
|||
4221 GGCTCTCCAGATGCTGGGGTTTACAGCGTGAAGCCACCGCCCAAGCCAAATGTC 4275
|||
3082 GGCTCTCCCAAGTGTCTGAGTTACAGGTGTGAAGCCACCGCACCTCTGTGATTC 3136
|||

| | | | |
|----|------|---|-------|
| Db | 5696 | CACCATGCGCTGGCCCTGCTGCTCTATGTTTGTATGTTGCCACAGAGTGTCTGAGTATCTCAC | 57555 |
| Qy | 671 | TGACTATGCTGATGTGATGACAGTACCTAAAGCCATCAGATC-----TACCCTT | 718 |
| Db | 5756 | TCCCTCCTCTGTGAGGCTCCCTTTGGTCAATTTCCCAAPACCTATTACAGTTTACCAAT | 5815 |
| Qy | 719 | TAAATATGCAGTTTGGGCGAGGACCGGTGCTCATGCTGTATATTCAGCACTTTGGAG | 778 |
| Db | 5816 | ATTAAAGATTGTGTGGGCGAGGCATGTGGCTCAAGCCTGTGTGTCCTCCGGCACTTTTGGAG | 58759 |
| Qy | 779 | GCAGAGGTGGGAGATCACTTGAGGCCAGAGTTTGACACAGCCTTGCCCAACATGTGGA | 838 |
| Db | 5876 | GCGGAGGCGGGGATGATCAGTAGGATCAGAGATTGAACACAGCTTGCCCAACATGTGTGA | 59359 |
| Qy | 839 | AACCTGTCTTTACTCTAAAAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTGCTGGGGC | 898 |
| Db | 5936 | AAACCCGCTCTACT-----AAAAAGTCAAAAATTTAGCTGGGGCTGTGTGGGAG | 59844 |
| Qy | 899 | ACACCTGTATCCAGCTATGCTGTGAGGCTGAGGACGAGATCACTTGAACCTGTGAGG | 958 |
| Db | 5985 | GCGCCTGTAGTCCAGCTGTCTCGGAGGCTGAGGACAGGAATTTCTTGAAGCCGGGAGG | 60444 |
| Qy | 959 | CGAGAGTTGCAGTGGGCGGAGATCAGATCACCGCCTTCCAGCTGTGGGAGACAGACAAAGA | 1018 |
| Db | 6045 | TAGAGGTTCAGGTGAGCCGAGACTGCTCATTTGCACTCCAGCCTGTGGGAGACAGAGTGA | 61040 |
| Qy | 1019 | CTCTGTCTCAATAATTAATTAATTAACAACGA | 1050 |
| Db | 6105 | CACGATCTCAAAATATATAACCAATTAAGA | 6136 |

| | |
|------------|---|
| RESULT 15 | |
| BC038630 | |
| LOCUS | |
| DEFINITION | BC038630 2772 bp mRNA linear HTC 04-MAR-2003 Homo sapiens, similar to hypothetical protein FLJ20489, clone IMAGE:5263792. mRNA. |
| ACCESSION | BC038630 |
| VERSION | BC038630.1 GI:24116283 |
| KEYWORDS | HTC. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (Bases 1 to 2772) Strauberg,R. TITLE Direct Submission JOURNAL Submitted (15-OCT-2002) National Institutes of Health, Mammalian |

| REMARK | COMMENT |
|--|------------------------|
| NIH-MGC Project URL: http://mgc.nci.nih.gov | Contact: MGC help desk |

CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shrikant Toshuyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadane@systemsbiology.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lindel.gov>
 Series: IRAC Plate: 73 Row: 4
 This clone has the following problem: retained intron.

| FEATURES | location/qualifiers |
|----------|--------------------------|
| SOURCE | 1..2772 |
| | /organism="Homo sapiens" |
| | /mol_type="mRNA" |
| | /db_xref="taxon:9606" |

```

/c1one="IMAGE:5263792"
/tissue_type="Brain/ hippocampus"
/c1one_lib="NIH_MGC_95"
/lab host="DH10B"
/note="Vector: pBluescript"

```

| | | |
|----|------------------------|---|
| | Query Match Similarity | 6.4%; Score 321.6; Length 2772; |
| | Best Local Similarity | 72.4%; Pred. No. 1.5e-18; |
| | Matches | 537; Conservative 0; Mismatches 174; Indels 31; Gaps 8; |
| OY | 327 | GNGGTCTATGATCATCTTCTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTC 386 |
| Dd | 2038 | GTGAATTACATAAACCCCTTTCTTTCTTTCTTTCTTTCTTTCTTTTGTAGATGSAATC 2097 |
| OY | 387 | TGCGCTGTGCGCCAAGCTGGAATGCAATGGCGGGAACTCGGCTCACGCAAGCTCCGC 446 |
| Dd | 2098 | TTCTCTGTGGCCAGGCT-GAGTGCAATGTGTGATCTCGGCTCACGCAACTCTAAC 2156 |
| OY | 447 | TCCCGGGTTCACGCATTCTCTCTCTCAGGCTCCCAAATGAGTGGGACTACAGGCGCC 506 |
| Dd | 2157 | TCTGTGGTTCAAGCAATTCTCTCTCTCAGGCTCCCAAATGAGTGGGACTACAGGCAATG 2216 |
| OY | 507 | GCCAATAAGCCCGGCTAATTTTTTTTGTATTTTAGTAGAGAGCGGGTTTACCCTTTAG 566 |
| Dd | 2217 | ACCACTACGCGCTGGCTGA--TTTTTGTATTTTGTAGAGGCAAGTTTACCAATGTTGG 2274 |
| OY | 567 | CCGGGATAGCCTGCATCTCTGACCTC--GTGATTCGCGCGGCTCGGCTCCCAAGTGC 624 |
| Dd | 2275 | TCAGGCTGG-TTTAACTCTGACCTCAGGAGCACTACCGGCTCGGCTTCAAGGTGT 2333 |
| OY | 625 | TGGATTACAAGCGCTGAGCCACCGCGCGGCGCATGATCATCTTCTGACTATGC----- 679 |
| Dd | 2334 | TGTATTACAGGATGAGCCACCACTCTGSCATATAACCTTTTCAGTAATTCATCTT 2393 |
| OY | 680 | -----TGATGTGACAACTAAGCATTAAGCATCACTTACCTTTAATATGAG--TTT 732 |
| Dd | 2394 | CTATTTAATTTAGCGAGGTTCACTTTAACACCAAGATAGTGAATATACCAAGAATCC 2453 |
| OY | 733 | GGGCGAGGACCGTGGGCTCATGCTGTAAATTCAGACACTTTGGGAGGAGAGTGGGTGA 792 |
| Dd | 2454 | TAGCTGGAGCTGTGGCTCATGCTGTAAATTCACCACTTTGGGAGGCTGAGGGAATGG 2513 |
| OY | 793 | ATCATTTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAATGATGTAAACTCTGTCTTAC 852 |
| Dd | 2514 | ATCATCTGAGGTCAGGAGTTGAGACCAAGCTGGCCAAATGATGTAAACTCTGTCTTAC 2564 |
| OY | 853 | TAAAAAAAAAAAAAAAAAAAAATCAAGCGGGGTGTCTGTGGGGCACACTGTAAATCCC 912 |
| Dd | 2565 | ---CAAAATATATCTAAATATACAGAAATTTGGCTGGGCGTGTGGGGCCACTGTGTCTCC 2621 |
| OY | 913 | AGCTATGCTGAGGCTGAGGACAGAGATCACTTGAACCCCTGGAGGGCGAGATTTGCAGTG 972 |
| Dd | 2622 | GGCTACTTTGGAGGCTGAGGACAGAGATCTGTGAACCCGGAGGATGCACTTTGCAAGC 2681 |
| OY | 973 | GGCGGAGATCAATCACCGCCCTTCAGCTGGGCGAC-AGAAGCAAGACTGTCTCAAT 1031 |
| Dd | 2682 | AGCGGAGATGACCAATTCGACCTCAAGCTGGGAGACAAAGCAAACTCGTCTTACCA 2741 |
| OY | 1032 | AATTAATTAACCAACGACCA 1053 |
| Dd | 2742 | AAAAAAAAAAAAAAAAAAAAA 2763 |

Search completed: February 27, 2005, 12:42:13
Job time : 15686.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:13 ; Search time 310.935 Seconds
(without alignments)
10530.153 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagagcccccaag.....tcgcggggggaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodate/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodate/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodate/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodate/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodate/1/ina/PTCTUS.COMB.seq: *
6: /cgn2_6/ptodate/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------------------|--------------------|
| 1 | 1999.4 | 99.9 | 10818 | 4 US-09-949-016-13583 | Sequence 13583, A |
| 2 | 601 | 30.0 | 601 | 4 US-09-949-016-63402 | Sequence 63402, A |
| 3 | 600.6 | 30.0 | 601 | 4 US-09-949-016-63399 | Sequence 63399, A |
| 4 | 600.6 | 30.0 | 601 | 4 US-09-949-016-63400 | Sequence 63400, A |
| 5 | 600.6 | 30.0 | 601 | 4 US-09-949-016-63401 | Sequence 63401, A |
| 6 | 600.6 | 30.0 | 601 | 4 US-09-949-016-63403 | Sequence 63403, A |
| 7 | 600.6 | 30.0 | 601 | 4 US-09-949-016-63404 | Sequence 63404, A |
| 8 | 599 | 29.9 | 601 | 4 US-09-949-016-63405 | Sequence 63405, A |
| 9 | 240.2 | 12.0 | 55827 | 3 US-09-813-133A-3 | Sequence 3, Appl1 |
| 10 | 240.2 | 12.0 | 55827 | 4 US-10-212-877-3 | Sequence 3, Appl1 |
| 11 | 235.6 | 11.8 | 75674 | 4 US-09-949-016-17557 | Sequence 17559, A |
| 12 | 233.2 | 11.7 | 265038 | 4 US-09-949-016-15779 | Sequence 15779, A |
| 13 | 229.8 | 11.5 | 131724 | 4 US-09-949-016-12893 | Sequence 12893, A |
| 14 | 226.2 | 11.3 | 109159 | 4 US-09-949-016-14169 | Sequence 14169, A |
| 15 | 226.2 | 11.3 | 109159 | 4 US-09-949-016-14169 | Sequence 14170, A |
| 16 | 225.6 | 11.3 | 18319 | 4 US-09-949-016-17446 | Sequence 17446, A |
| 17 | 225.4 | 11.3 | 325791 | 4 US-09-768-185A-1 | Sequence 1, Appl1 |
| 18 | 225.2 | 11.3 | 601 | 4 US-09-949-016-202742 | Sequence 202742, A |
| 19 | 224.6 | 11.2 | 72992 | 4 US-09-949-016-17552 | Sequence 17592, A |
| 20 | 221.4 | 11.1 | 70131 | 4 US-09-949-016-16064 | Sequence 16064, A |
| 21 | 221.2 | 11.1 | 150394 | 4 US-09-949-016-13042 | Sequence 13042, A |
| 22 | 220.8 | 11.0 | 209210 | 4 US-09-949-016-15094 | Sequence 15094, A |
| 23 | 220 | 11.0 | 136264 | 4 US-09-949-016-12756 | Sequence 12756, A |
| 24 | 220 | 11.0 | 136265 | 4 US-09-949-016-13001 | Sequence 13001, A |
| 25 | 219.8 | 11.0 | 166698 | 4 US-09-949-016-16038 | Sequence 16038, A |
| 26 | 219.4 | 11.0 | 35058 | 4 US-09-949-016-12607 | Sequence 12607, A |
| 27 | 219.4 | 11.0 | 35059 | 4 US-09-949-016-13831 | Sequence 13831, A |

| | | | | | |
|----|-------|------|--------|------------------------|-------------------|
| 28 | 219.4 | 11.0 | 165651 | 4 US-09-949-016-13032 | Sequence 13032, A |
| 29 | 219.2 | 11.0 | 16984 | 4 US-09-949-016-16893 | Sequence 16893, A |
| 30 | 218.6 | 10.9 | 17353 | 4 US-09-949-016-13437 | Sequence 13437, A |
| 31 | 218.6 | 10.9 | 301828 | 4 US-09-949-016-13969 | Sequence 13969, A |
| 32 | 218.4 | 10.9 | 15358 | 4 US-09-949-016-13382 | Sequence 13382, A |
| 33 | 218.2 | 10.9 | 91831 | 4 US-09-949-016-13694 | Sequence 13694, A |
| 34 | 218 | 10.9 | 88490 | 4 US-09-949-016-12758 | Sequence 12758, A |
| 35 | 218 | 10.9 | 88766 | 4 US-09-949-016-14232 | Sequence 14222, A |
| 36 | 217.8 | 10.9 | 601 | 4 US-09-949-016-158456 | Sequence 58456, A |
| 37 | 217.8 | 10.9 | 30371 | 4 US-09-949-016-15395 | Sequence 15395, A |
| 38 | 217.4 | 10.9 | 44789 | 4 US-09-949-016-13909 | Sequence 13909, A |
| 39 | 217.2 | 10.9 | 47751 | 4 US-09-949-016-12812 | Sequence 12812, A |
| 40 | 217.2 | 10.9 | 47752 | 4 US-09-949-016-16955 | Sequence 16955, A |
| 41 | 216.8 | 10.8 | 107421 | 4 US-09-949-016-15532 | Sequence 15532, A |
| 42 | 216.4 | 10.8 | 99797 | 4 US-09-949-016-15255 | Sequence 15255, A |
| 43 | 216.4 | 10.8 | 256171 | 4 US-09-949-016-12822 | Sequence 12822, A |
| 44 | 216.4 | 10.8 | 256176 | 4 US-09-949-016-15524 | Sequence 15524, A |
| 45 | 216.2 | 10.8 | 8230 | 4 US-09-949-016-14183 | Sequence 14183, A |

ALIGNMENTS

```
RESULT 1
US-09-949-016-13583
; Sequence 13583, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13583
; LENGTH: 10818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13583

Query Match          99.9%; Score 1999.4; DB 4; Length 10818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GACACCAAGAGCCCCCAAGAGTGTCCAGGTTCTTGTATACGACCAATCTCTGC 60
        |||
DB      3245 GACACCAAGAGCCCCCAAGAGTGTCCAGGTTCTTGTATACGACCAATCTCTGC 3304

QY      61 CGCCTTCATCTTCTTCACATTTCTCATTTGTTCTCTTTGACAGTGCATCTCCTAAG 120
        |||
DB      3305 CGCCTTCATCTTCTTCACATTTCTCATTTGTTCTCTTTGACAGTGCATCTCCTAAG 3364

QY      121 TCTCCAGGCTGGAATATCAGGCTGCTAAGTCAGAAAAGTGCAGAGATCTTACCCGA 180
        |||
DB      3365 TCTCCAGGCTGGAATATCAGGCTGCTAAGTCAGAAAAGTGCAGAGATCTTACCCGA 3424

QY      181 GACAGATAGATGACACCATTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGCGCAG 240
        |||
DB      3425 GACAGATAGATGACACCATTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGCGCAG 3484

QY      241 TGAGACACTCTCTTTATTCAGACATACACTGAGTGCACACTCGGTAATGAGAGC 300
        |||
DB      3485 TGAGACACTCTCTTTATTCAGACATACACTGAGTGCACACTCGGTAATGAGAGC 3544

QY      301 GTTGCCAAATTCTGAGATTCAGCAATTCGCAAGAGTCCAGACCCCTGTTCTCAGAG 360
```

3545 GTTGCAGAAATTCGAGAAATCCAGCAATTCGCAAGACGTGAGACCCCTGTTCTCACAGA 3604
361 GCTCATACCTTAAGATAGTGTGTTTATAGTAATAATAGTGTGCTGTTATATCATTTTC 420
3605 GCTCATACCTTAAGATAGTGTGTTTATAGTAATAATAGTGTGCTGTTATATCATTTTC 3664
421 CAGTTTTTATAGTACCATTTAAACAGGTAAAGAGCTGGCGCAGTGGCTCACACCT 480
3665 CAGTTTTTATAGTACCATTTAAACAGGTAAAGAGCTGGCGCAGTGGCTCACACCT 3724
481 GTATCCACAGACTTTTGGAGGCTGAGGCGAGATCACCTTTGTCAGAGCTTTGAGA 540
3725 GTATCCACAGACTTTTGGAGGCTGAGGCGAGATCACCTTTGTCAGAGCTTTGAGA 3784
541 CTGAGCTGGCCAAATGCGGAAATCTGTCCTTAAATAAATAATGAGCTTGGCCTGGC 600
3785 CTGAGCTGGCCAAATGCGGAAATCTGTCCTTAAATAAATAATGAGCTTGGCCTGGC 3844
601 ATGATGGCGGCGCTGTATCTCAGCTGCTCAGAGCGGAGACACAAAGATCACTTAA 660
3845 ATGATGGCGGCGCTGTATCTCAGCTGCTCAGAGCGGAGACACAAAGATCACTTAA 3904
661 ACCAGAGAGGTGAGGTTGCAGTGAGCTGAGATCTGTCCTCACTTCCAACTGGAGAC 720
3905 ACCAGAGAGGTGAGGTTGCAGTGAGCTGAGATCTGTCCTCACTTCCAACTGGAGAC 3964
721 AGAGTGACCTTTTGTCTCAAAAAGAAAAAACAATTAATAAAGAAACAGGTGAAG 780
3965 AGAGTGACCTTTTGTCTCAAAAAGAAAAAACAATTAATAAAGAAACAGGTGAAG 4024
781 TTAACTTATTAACCAATGTATCCAAATACATCATTTCAAGTGTATTAATAA 840
4025 TTAACTTATTAACCAATGTATCCAAATACATCATTTCAAGTGTATTAATAA 4084
841 ACAATTATGATAGATTAATTACATTTCTTTCTTTTCAATTAAGCTTTTGAAGT 900
4085 ACAATTATGATAGATTAATTACATTTCTTTCTTTTCAATTAAGCTTTTGAAGT 4144
901 GAGTATATATGATATGCTGACAGACATCTCAATTTGGAATGATCTTCACTTCAAGGTGC 960
4145 GAGTATATATGATATGCTGACAGACATCTCAATTTGGAATGATCTTCACTTCAAGGTGC 4204
961 AGTAGCACATGTGCTAGCAGTACTGTATGATGATGAGCAGGATCTAGAGGAAAGATC 1020
4205 AGTAGCACATGTGCTAGCAGTACTGTATGATGATGAGCAGGATCTAGAGGAAAGATC 4264
1021 AGGGCTGTTTTGTATGTTGGGCAAGTTGTGCACTGCATTAAGATACATATCTAATAG 1080
4265 AGGGCTGTTTTGTATGTTGGGCAAGTTGTGCACTGCATTAAGATACATATCTAATAG 4324
1081 GGCACCTCCGTGTACAGATGTCAAGTTTGGCACTTTTCAAGGCTGGTGTAAAGTGC 1140
4325 GGCACCTCCGTGTACAGATGTCAAGTTTGGCACTTTTCAAGGCTGGTGTAAAGTGC 4384
1141 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTCTGGTAAATATCTTGA 1200
4385 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTCTGGTAAATATCTTGA 4444
1201 GGAAGAAAAAGAAATCTGTGATGATTTTCAAGAGAAATTTTAAATACAGGGGATTA 1260
4445 GGAAGAAAAAGAAATCTGTGATGATTTTCAAGAGAAATTTTAAATACAGGGGATTA 4504
1261 ATTGCAAAAGCTGTGAGAGGCTGAGAGAAACAAGTTAAATAAATAAATAAATAAATAA 1320
4505 ATTGCAAAAGCTGTGAGAGGCTGAGAGAAACAAGTTAAATAAATAAATAAATAAATAA 4564
1321 AAGAACTGATTAATAGGCAATTTCAAGAGTGTGTAAGTTAAATCCCAAAATATAAC 1380
4565 AAGAACTGATTAATAGGCAATTTCAAGAGTGTGTAAGTTAAATCCCAAAATATAAC 4624
QY ATGTTTTTATGATTAATTAAGGCAATTAAGGCAATTAATTAAGGCAATTAATTAAGGCA 1440
ATGTTTTTATGATTAATTAAGGCAATTAAGGCAATTAATTAAGGCAATTAATTAAGGCA

4625 ATGTTTTTATGATTAATTAAGGCAATTAAGGCAATTAATTAAGGCAATTAATTAAGGCA 4684
QY 1441 CTGAGAGAGTGGCAATTTGAGAGAGAAATGATGACACAAAGAGTAAATCTGTGAAGT 1500
4685 CTGAGAGAGTGGCAATTTGAGAGAGAAATGATGATGACACAAAGAGTAAATCTGTGAAGT 4744
QY 1501 TTAAAGGGAAGAAAGAGCAGTGCAGAGGCTGAGGAGTAAAGAAATTTGGCTGATTC 1560
4745 TTAAAGGGAAGAAAGAGCAGTGCAGAGGCTGAGGAGTAAAGAAATTTGGCTGATTC 4804
QY 1561 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAGTGGGGCAACATGTAAGAGT 1620
4805 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAGTGGGGCAACATGTAAGAGT 4864
QY 1621 AGGCTGAGAGTGAAGGCAAGGCGAATGCTGCAAGTATTTCTTGGTCAACACAGAG 1680
4865 AGGCTGAGAGTGAAGGCAAGGCGAATGCTGCAAGTATTTCTTGGTCAACACAGAG 4924
QY 1681 GCTTCCCTATGTTCTAATGGAAGTGTATCTGTTGAGGAAGACAGAAATTAATCAAC 1740
4925 GCTTCCCTATGTTCTAATGGAAGTGTATCTGTTGAGGAAGACAGAAATTAATCAAC 4984
QY 1741 TGTATCATTAACACAGACCTTTCTGTATTAAGGCTCCCAAGGATCTTAAGAGACGTA 1800
4985 TGTATCATTAACACAGACCTTTCTGTATTAAGGCTCCCAAGGATCTTAAGAGACGTA 5044
QY 1801 AGTTAACAAAGCTCTCATTAAGGAGGAGTGTGTTTCAAGTATTAAGAAAGTGGGAT 1860
5045 AGTTAACAAAGCTCTCATTAAGGAGGAGTGTGTTTCAAGTATTAAGAAAGTGGGAT 5104
QY 1861 CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCAATGTAACAACTTGACAA 1920
5105 CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCAATGTAACAACTTGACAA 5164
QY 1921 ATGCAACACTCCCAAGCTCTCCCAATTTCTGTCGTGCTGAGTGTGGGGGGTGG 1980
5165 ATGCAACACTCCCAAGCTCTCCCAATTTCTGTCGTGCTGAGTGTGGGGGGTGG 5224
QY 1981 TTGCGAGGGGGAACCTTTTA 2001
5225 TTGCGAGGGGGAACCTTTTA 5245
Db 5225 TTGCGAGGGGGAACCTTTTA 5245

RESULT 2
US-09-949-016-63402/C
; Sequence 63402, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63402
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63402

Query Match 30.0%; Score 601; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.7e-147;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 AGGCTGGCGCAGTGGCTTCACTGTAATCCAGCACTTTGGAGGCTGAGGCAAGGAG 515

Db 601 AGGCTGGGCGGAGTGGCTACACCTGTATCCAGACACTTTGGAGGTGAGGAGGAG 542
Qy 516 ATCACTTTGGTCAAGAGTTTGAAGTCTAGCTGGCCAACTGGGAACTCTGTCTTAA 575
Db 541 ATCACTTTGGTCAAGAGTTTGAAGTCTAGCTGGCCAACTGGGAACTCTGTCTTAA 482
Qy 576 AAAAAATPACAAAATTAAGCTGGCATGTGGCCGGGCGCTGTATCTCAGCTGCTGAG 635
Db 481 AAAAAATPACAAAATTAAGCTGGCATGTGGCCGGGCGCTGTATCTCAGCTGCTGAG 422
Qy 636 AGCCGAGACACAAAGATCACTTAAACCCAGAGAGTGGAGGTGCTGAGTGAATCG 695
Db 421 AGCCGAGACACAAAGATCACTTAAACCCAGAGAGTGGAGGTGCTGAGTGAATCG 362
Qy 696 TGCCACTCACTCCAACTGGGAGACAGAGTGACCTTTTGTCTCAAAAAAGAAAAA 755
Db 361 TGCCACTCACTCCAACTGGGAGACAGAGTGACCTTTTGTCTCAAAAAAGAAAAA 302
Qy 756 ACAAGTAAAAAGAACAGGTGAAGTTAACTTTAAACCAATGTATCCCAATACAA 815
Db 301 ACAAGTAAAAAGAACAGGTGAAGTTAACTTTAAACCAATGTATCCCAATACAA 242
Qy 816 CATTTCAAGTGTATTAATTAATAACATTAATGAGATGACTTACATTTCTTTCTT 875
Db 241 CATTTCAAGTGTATTAATTAATAACATTAATGAGATGACTTACATTTCTTTCTT 182
Qy 876 GTTTTCATTAATTAAGTCTTTGAAAGTGAATATATATGTATGCTGACAGCATCTCA 935
Db 181 GTTTTCATTAATTAAGTCTTTGAAAGTGAATATATATGTATGCTGACAGCATCTCA 122
Qy 936 TGdACTGACTCACTTTCAGAGTCTCAGTACCAATGTGGTACACTTACATTTCTG 995
Db 121 TGdACTGACTCACTTTCAGAGTCTCAGTACCAATGTGGTACACTTACATTTCTG 62
Qy 996 TGdGACGAGTCTGAGGAAAGATCAGGCTGTTTGTATGCTTGAGGAGGTTGCACT 1055
Db 61 TGdGACGAGTCTGAGGAAAGATCAGGCTGTTTGTATGCTTGAGGAGGTTGCACT 2
Qy 1056 G 1056
Db 1 G 1

RESULT 3
US-09-949-016-63399/c
; Sequence 63399, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63399
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63399

Query Match 30.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 148 TAAGTCAGAAAAGGTGAGAGTGTCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 207
Db 601 TAAGTCAGAAAAGGTGAGAGTGTCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 542
Qy 208 CCGGTGACAAAGAGGTAGAGTCTCTGCGAGGTGAGGACACTCTTTTATTCAGCAGA 267
Db 541 CCGGTGACAAAGAGGTAGAGTCTCTGCGAGGTGAGGACACTCTTTTATTCAGCAGA 482
Qy 268 TACACACTGAGTCCCACTGCTGTAACATGAGAGGTTGCCAAATTTGAGAAATCCAGCA 327
Db 481 TACACACTGAGTCCCACTGCTGTAACATGAGAGGTTGCCAAATTTGAGAAATCCAGCA 422
Qy 328 TGCCAAAGACAGTACGAGACCCCTGTTCTACAGAGCTCATACCTAGAGTGTGTGTTA 387
Db 421 TGCCAAAGACAGTACGAGACCCCTGTTCTACAGAGCTCATACCTAGAGTGTGTGTTA 362
Qy 388 GTAGAAATTAATGCTGAGTCTTATGTATGATTTTCAGTTTATAGTCCATTAACA 447
Db 361 GTAGAAATTAATGCTGAGTCTTATGTATGATTTTCAGTTTATAGTCCATTAACA 302
Qy 448 GGTAAAAAGGCTGGGCGCAGTGGCTCAACCTGTATATCCAGACTTTGGAGGCTGAG 507
Db 301 GGTAAAAAGGCTGGGCGCAGTGGCTCAACCTGTATATCCAGACTTTGGAGGCTGAG 242
Qy 508 GCAGGAGATCACCTTTGCTGAGAGTTGAGACTGAGTGGCCATGCGCAATGCGCAACTCT 567
Db 241 GCAGGAGATCACCTTTGCTGAGAGTTGAGACTGAGTGGCCATGCGCAATGCGCAACTCT 182
Qy 568 GTCTTAAAAAAAATTAACAAAATTAAGCTGAGTGTGGGCGGCTGTAAATCTCAGC 627
Db 181 GTCTTAAAAAAAATTAACAAAATTAAGCTGAGTGTGGGCGGCTGTAAATCTCAGC 122
Qy 628 TGCTCAGAGGCGGAGACACAGATCACTTAAACCCAGAGGTGAGAGTTGAGTGAAC 687
Db 121 TGCTCAGAGGCGGAGACACAGATCACTTAAACCCAGAGGTGAGAGTTGAGTGAAC 62
Qy 688 TGAGATGCTGCTCACTCACTGAGGAGAGAGTGAACATTTTGTCTCAAAAAAGAA 747
Db 61 TGAGATGCTGCTCACTCACTGAGGAGAGAGTGAACATTTTGTCTCAAAAAAGAA 2
Qy 748 A 748
Db 1 A 1

RESULT 4
US-09-949-016-63400/c
; Sequence 63400, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63400
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63400

Query Match 30.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| OY | 167 | GGATCTCTAACCCGAGACAGATAGATGACACCAATGTTCTGCGCCGGGAGCAAAAGCAGTGG | 226 |
| Db | 601 | GGATCTCTTACCAGAGACAGATAGATGACACCAATGTTCTGCGCCGGGAGCAAAAGCAGTGG | 542 |
| OY | 227 | AGACTCTCGCAGGAGTGAGGACACTCTCTTTATTCAGCAGATACACATGAGTGGCAACT | 286 |
| Db | 541 | AGACTCTCGCAGGAGTGAGGACACTCTCTTTATTCAGCAGATACACATGAGTGGCAACT | 482 |
| OY | 287 | CGGTAACTAGAGGCGTGGCCAAATCTGGAATCCAGCAATTGGCCAAAGACATGACAGAAC | 346 |
| Db | 481 | CGGTAACTAGAGGCGTGGCCAAATCTGGAATCCAGCAATTGGCCAAAGACATGACAGAAC | 422 |
| OY | 347 | CCTGTTCTCACAGAGCTCATACCTTAGAGTAGTGGTGTTTAGTAAGAAATATATGCTAGCT | 406 |
| Db | 421 | CCTGTTCTCACAGAGCTCATACCTTAGAGTAGTGGTGTTTAGTAAGAAATATATGCTAGCT | 362 |
| OY | 407 | GCTTATGTCAATTTCCAGTTTCTTTAGTAGCCACATTAAACAGGTAAAAAAAGCTGGGCGC | 466 |
| Db | 361 | GCTTATGTCAATTTCCAGTTTCTTTAGTAGCCACATTAAACAGGTAAAAAAAGCTGGGCGC | 302 |
| OY | 467 | AGTGCTCACACCTGTAAATCCAGACCTTTGGAGGCTGAGGAGGCAATACCTTTGG | 526 |
| Db | 301 | AGTGCTCACACCTGTAAATCCAGACCTTTGGAGGCTGAGGAGGCAATACCTTTGG | 242 |
| OY | 527 | TCAGAGTTTGAAGACTAGGCTTGCCCAACATGCGCAAACTCTGTCCTTAAAAAAAAATACA | 586 |
| Db | 241 | TCAGAGTTTGAAGACTAGGCTTGCCCAACATGCGCAAACTCTGTCCTTAAAAAAAAATACA | 182 |
| OY | 587 | AAAAATTAGCTTGCAATGTGTGGCGGGCGCTGTAAATCTCAGCTGCTCAGAGGCGCAGACA | 646 |
| Db | 181 | AAAAATTAGCTTGCAATGTGTGGCGGGCGCTGTAAATCTCAGCTGCTCAGAGGCGCAGACA | 122 |
| OY | 647 | CAAGATCACTTAAACCCAGAGGTGAGAGTTGACATGAGCTGAGATCGTGCCACTCACT | 706 |
| Db | 121 | CAAGATCACTTAAACCCAGAGGTGAGAGTTGACATGAGCTGAGATCGTGCCACTCACT | 62 |
| OY | 707 | CCAACTTGAGAGACAGATGACATTTTGTCTCAAAAAAAGAAAAAACAAGTAAAAA | 766 |
| Db | 61 | CCAACTTGAGAGACAGATGACATTTTGTCTCAAAAAAAGAAAAAACAAGTAAAAA | 2 |
| OY | 767 | A 767 | |
| Db | 1 | A 1 | |

```

RESULT 5
US-09-949-016-63401/C
; Sequence 63401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63401
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63401

Query Match: 30.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;

```

| | Matches | 600; | Conservative | 1; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|----|--------|----|------|----|
| QY | 242 | GAGGACACCTCTCTTTATTTCAGCAGATACACATGAGTGGCCAACTCGGTAAACAATGAGGG | 301 | | | | | | | |
| Db | 601 | GAGGACACCTCTCTTTATTTCAGCAGATACACATGAGTGGCCAACTCGGTAAACATGAGGG | 542 | | | | | | | |
| QY | 302 | TTGCGCAATTCAGAGATCCAGCAATTCSCAAGACATGACGAGCCCGCTGTTCTCACAGAG | 361 | | | | | | | |
| Db | 541 | TTGCGCAATTCAGAGATCCAGCAATTCGCCAAGACATGACGAGCCCGCTGTTCTCACAGAG | 482 | | | | | | | |
| QY | 362 | CTCATACCTTAGAGTAGTGTGTTTAGTAGAAATATATGCTGAGCTGCTTATGTCAATTC | 421 | | | | | | | |
| Db | 481 | CTCATACCTTAGAGTAGTGTGTTTAGTAGAAATATATGCTGAGCTGCTTATGTCAATTC | 422 | | | | | | | |
| QY | 422 | AGTTTTTTAGTAGCCACATTAATAACAGGTAAATAAAAGCTGGGCGCAGTGGCTCACACCTG | 481 | | | | | | | |
| Db | 421 | AGTTTTTTAGTAGCCACATTAATAACAGGTAAATAAAAGCTGGGCGCAGTGGCTCACACCTG | 362 | | | | | | | |
| QY | 482 | TAATCCAGCAGCTTGGGAGGCTGAGGACGAGCAGATCACTTTGGTCAAGAGTTTGAGAC | 541 | | | | | | | |
| Db | 361 | TAATCCAGCAGCTTGGGAGGCTGAGGACGAGCAGATCACTTTGGTCAAGAGTTTGAGAC | 302 | | | | | | | |
| QY | 542 | TAGCTGGCCCAACATGGGAAACTGTGTCTTAATAAAAAATTAACAAAATTATAGCTGGCA | 601 | | | | | | | |
| Db | 301 | YAGCTGGCCCAACATGGGAAACTGTGTCTTAATAAAAAATTAACAAAATTATAGCTGGCA | 242 | | | | | | | |
| QY | 602 | TGTTGTCGGGCGCCTGTATCTCAGCTGCTCAGGAGCGCAGACACAAGAAATCACTTAA | 661 | | | | | | | |
| Db | 241 | TGTTGTCGGGCGCCTGTATCTCAGCTGCTCAGGAGCGCAGACACAAGAAATCACTTAA | 182 | | | | | | | |
| QY | 662 | CCGAGAGGTGAGAGTTGCACTGAGTGAATTCGTGCCATCTCCAACTCGGAGAC | 721 | | | | | | | |
| Db | 181 | CCGAGAGGTGAGAGTTGCACTGAGTGAATTCGTGCCATCTCCAACTCGGAGAC | 122 | | | | | | | |
| QY | 722 | GAGTGAACCTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAGT | 781 | | | | | | | |
| Db | 121 | GAGTGAACCTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAGT | 62 | | | | | | | |
| QY | 782 | TAACTTTAATTAACCAATGTATCCCAATACATCATTTCAAAGTGAATTATATTAATA | 841 | | | | | | | |
| Db | 61 | TAACTTTAATTAACCAATGTATCCCAATATCAATCATTTCAAAGTGAATTATATTAATA | 2 | | | | | | | |
| QY | 842 | C 842 | | | | | | | | |
| Db | 1 | C 1 | | | | | | | | |

```

        6 RESULT
        6 US-09-949-016-63403/c
        6 Sequence 63403, Application US/09949016
        6 Patent No. 6812339
        6 GENERAL INFORMATION:
        6 APPLICANT: VENTER, J. Craig et al.
        6 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
        6 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
        6 FILE REFERENCE: CLO01307
        6 CURRENT APPLICATION NUMBER: US/09/949,016
        6 CURRENT FILING DATE: 2000-04-14
        6 PRIOR APPLICATION NUMBER: 60/241,755
        6 PRIOR FILING DATE: 2000-10-20
        6 PRIOR APPLICATION NUMBER: 60/237,768
        6 PRIOR FILING DATE: 2000-10-03
        6 PRIOR APPLICATION NUMBER: 60/231,498
        6 PRIOR FILING DATE: 2000-09-08
        6 NUMBER OF SEQ ID NOS: 207012
        6 SOFTWARE: FASTSBQ For Windows Version 4.0
        6 SEQ ID NO 63403
        6 LENGTH: 601
        6 TYPE: DNA
        6 ORGANISM: Human
        6 US-09-949-016-63403
        6 Query Match
        6 30.0%; Score 600.6; DB 4; Length 601;

```


Query Match 29.9%; Score 599; DB 4; Length 601;
Best Local Similarity 99.7%; Pred. No. 5.6e-147;
Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1301 AAAAAAAAACTGTGTGTGTCAGATCTGCATTAATATGAGCAATTCAGAGAGTGTAA 1360
    |||
Db 601 AAAAAAAAACTGTGTGTGTCAGATCTGCATTAATATGAGCAATTCAGAGAGTGTAA 542

QY 1361 GGTAAACCCCAATTAATCAATGTTTATGATATGTAACATTAAGGCCAATTTCAA 1420
    |||
Db 541 GGTAAACCCCAATTAATCAATGTTTATGATATGTAACATTAAGGCCAATTTCAA 482

QY 1421 AAGTGTGTCAGAGGAGCTCTTGGAGAGGTGGCATTTGAGCAGAGATGATGACAA 1480
    |||
Db 481 AAGTGTGTCAGAGGAGCTCTTGGAGAGGTGGCATTTGAGCAGAGATGATGACAA 422

QY 1481 AGAAGCTAACTGTGAAGTTTAAAGGGAAAAAGAAAGGCACTGCAAAAGCCCTGAGCA 1540
    |||
Db 421 AGAAGCTAACTGTGAAGTTTAAAGGGAAAAAGAAAGGCACTGCAAAAGCCCTGAGCA 362

QY 1541 GTAAGAAATTTGCTGTATTTCAAAAGAAAGAAAGCAATGCACTGGAACAAAAGT 1600
    |||
Db 361 GTAAGAAATTTGCTGTATTTCAAAAGAAAGAAAGCAATGCACTGGAACAAAAGT 302

QY 1601 GGGGAGCAACAGTAAAGTGAAGTGAAGGTGAGGAGGGGGAATGCTGTGCAAGTAT 1660
    |||
Db 301 GGGGAGCAACAGTAAAGTGAAGTGAAGGTGAGGAGGGGGAATGCTGTGCAAGTAT 242

QY 1661 TTCTTGTCACCAACACAGAGCTTCCCTATGTTCTAATGAGCTGTATCTGTAGGAA 1720
    |||
Db 241 TTCTTGTCACCAACACAGAGCTTCCCTATGTTCTAATGAGCTGTATCTGTAGGAA 182

QY 1721 GAAAGAAATTAATCAAACTGTATCAACCAAGCACTTCTCTGTATTCAGGCTCCC 1780
    |||
Db 181 GAAAGAAATTAATCAAACTGTATCAACCAAGCACTTCTCTGTATTCAGGCTCCC 122

QY 1781 AAGGAGCTGAGAGAGAGTAACTTAACAAGCTCTCATTTAGCAGAGTGTGTTCACAG 1840
    |||
Db 121 AAGGAGCTGAGAGAGAGTAACTTAACAAGCTCTCATTTAGCAGAGTGTGTTCACAG 62

QY 1841 TAGTTAGAAAGTGGGAGTTCAGAGTACTCCAGTCCCATGAGTATGAAAAAGTCCCCC 1900
    |||
Db 61 TAGTTAGAAAGTGGGAGTTCAGAGTACTCCAGTCCCATGAGTATGAAAAAGTCCCCC 2

QY 1901 A 1901
    |||
Db 1 A 1
```

RESULT 9

US-09-813-133A-3/c
; Sequence 3, Application US/09813133A
; Patent No. 6455294
; GENERAL INFORMATION:
; APPLICANT: GAN, weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001173
; CURRENT APPLICATION NUMBER: US/09/813,133A
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Human
US-09-813-133A-3

Query Match 12.0%; Score 240.2; DB 3; Length 55827;
Best Local Similarity 68.6%; Pred. No. 4e-52;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

```
QY 363 TCATACCTAGAGTAGTGGTGTAGTAGAATTAATGCTGAGCTGCTTATGCTTCCA 422
    |||
Db 46943 TGAATCTGACACAGAGTGTGTATTAAGAAATATATGCACTGACATATTACTTTAA 46884

QY 423 GTTTTATTAGCCACATTAATAACAGGTAAAA--AAGCTGGGCGCAGTGGCTCACACCT 480
    |||
Db 46883 ATTTTCAAGTTACACCTTTTAAGAAAAATTAAGAAAGAGGCAAGGCAAGTGGCTCACACCT 46824

QY 481 GTATTCACAGCACTTTGGAGGCTGAGGAGGCAAGTCACTTTGGTCAAGAGTTTGAGA 540
    |||
Db 46823 ATATTCACAGCACTTTGGAGGCTGAGGAGGCGGATCACTTGAGATCGAGAGTTTGAGA 46764

QY 541 CTAGCTGGCCAAACATGGGAAACTGTCTGTAAAAAAAATACAAAATTAGCTGGC 600
    |||
Db 46763 CAGCAGAGCCAAACATGGGAAACCCTGTCTTACCAAAA--TACAAAATATGCCAGGC 46706

QY 601 ATGTGTGGCGGCGCTGTATCTCAGCTGCTCAGAGGCCGAGACCAAGAACTCTTAA 660
    |||
Db 46705 ATATGTGTGCAAGCCTGTATCTCAGCTGCTCAGAGGCCGAGACCAAGAACTCTTGA 46646

QY 661 ACCGAGAGGTGAGGTGCTGAGTGAAGTCAATCGTCCACT--CACTCCAACTGGGAGA 719
    |||
Db 46645 ACCGAGAGGCAAGAGGTGCTGAGTGAAGTCAATCGTCACTCCAGCCTGGATGA 46586

QY 720 CAGAGTGACA--CTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGA 778
    |||
Db 46585 CAGACAAACACCTGTCACACACACACACACAAAAAGAAATATGTGATTAATTAATA 46526

QY 779 AGTTAATTAAATTAACCAATGATATCCAAATACATCTTCAAGTAAATTAATATA 838
    |||
Db 46525 TTGTATTAATTAATTAACCAATATGTACAAAGTATCA---TTTCAAGTGAACCAATATA 46470

QY 839 AAACAATTATGATGAGTACTTTACATCTTTCTGTTTCAATTAAGTCTTGAAA 898
    |||
Db 46469 AA--AATTAATCTTGAATTAATTAATTAATTAAGTACTT-----TTTCAATCTAAGCCTTCAAAAT 46417

QY 899 GTGAGATATATGTTATGCTGACAGACATCTCAATTTGAGCTAGTCAATTTACAGGTGC 958
    |||
Db 46416 TCTGGTGTATTAATTAACACTGAGTGTATGTCTCAATTCAGACAGACTTAATTAAGCAC 46357

QY 959 TCAGTAGCCCACTGTGGCTAGCAATTAATGATTTGA 995
    |||
Db 46356 TGAATATGCAACATGTGTGTTAGTGGCTACTATGCTGA 46320
```

RESULT 10

US-10-212-877-3/c
; Sequence 3, Application US/10212877
; Patent No. 6818429
; GENERAL INFORMATION:
; APPLICANT: GAN, weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001173DIV
; CURRENT APPLICATION NUMBER: US/10/212,877
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/813,133
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-877-3

Query Match 12.0%; Score 240.2; DB 4; Length 55827;
Best Local Similarity 68.6%; Pred. No. 4e-52;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

Db 46943 TGAATCTTCAGACACGAGGTTGTTTAAATAGAAATATATATGCAAGTCAATATTTACTTTAA 46884
 Oy 423 GTTTTAAAGTACCCACATTAAAAACAGTAAAA--AAGCTGGGCGGAGTGAGCTCACACCT 480
 Db 46883 ATTTTCAAGTTACCACTTTAAAGAAAATAAAAAGGCGACGAGCTGCTACACCT 46824
 Oy 481 GTAAATCCACGACCTTTGGGAGCTGAGCGAGCAAGATCACTTTGGTCAGGAATTTGAGA 540
 Db 46823 AATAATCCACGACCTTTGGGAGCTGAGGAGCGGAGATCACTTGAAGATCAGGAGTTTGAC 46764
 Oy 541 CTAGCTGGCCAAACATGGCGAAACTGTCTCTAAAAAAAAAATAAAAAATTAAGCTGGC 600
 Db 46763 CCGAGGAGGCCCAACATGATGTAAACCCGCTCTTACCAAAAA--TACAAAATATGCCAGGC 46706
 Oy 601 ATGTGGCGGGGCGCCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAGAACTCACTTAA 660
 Db 46705 ATAGTGTGTACAGCCTGTATCTCCAGCTACTTGGGAGGCTGAGGACAGAACTTGCTTGA 46646
 Oy 661 ACCGAGGAGTGAGAGTTTGCATGAGCTGAGATCTGCGCACT--CATTCACACTTGGAGA 719
 Db 46645 ACCGAGGAGGCAAGAGTTTCAGTACGTAGACATGACATGCACTCCAGCTCCAGCTGAGTGA 46586
 Oy 720 CAGAGTGACA--CTTTTGTCTCAAAAAAGAAAAAAAACAAGTAAAAAAGAAACAGGTGA 778
 Db 46585 CAGAACAGACCCCTGCGACACACACACACAAAAAATATGTAGATTAATTCAAAT 46526
 Oy 779 AGTTAACTTTAAATPAAACCAATGTATCCAAATACAAATCATTTCAAAAGTAAATTAATATA 838
 Db 46525 TTGTATTTATTTAAACCAATATGTACAAAGTATCA----TTTCAAGTGAACCAATATA 46470
 Oy 839 AAACAATATAGAAATAGATTAATTAATCTTTCTTGTGTTTATATATTAAGCTTTGAAA 898
 Db 46469 AA--AATATCTTTGAAATATTTAATTAATTAATCTT-----TTTATATCTAAGCCTTCAAT 46417
 Oy 899 GTGATATATATGTATATGCTGACAGACACATCTCAATTTGGACTAGTACATTTACAGTGC 958
 Db 46416 TCTGTGTGTATTTAAACATCACTGATGTATGTCTCAATTCAGACAGACGTACTTAAAGCAC 46357
 Oy 959 TCAGTAGCCACATGTGGCTAGACAGTTATCTGTATTCGA 995
 Db 46356 TGGATAGCCACATGTGGTTAGTGGCTACTATGCTGGA 46320

RESULT 11
 US-09-949-016-17597/C
 ; Sequence 17597, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 17597
 ; LENGTH: 75674
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(75674)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17597

```

Query Match      11.8%; Score 235.6; DB 4; Length 75674;
Best Local Similarity 67.4%; Pred. No. 7,3e-51;
Matches 407; Conservative 0; Mismatches 189; Indels 8; Gaps 5;

QY      362 CTCATACCTTACAGAGTAGTGAGTGTATTAAGTAGAAT--AATGCTAGAGCTGTTATGTCAATT 419
DB      64429 CTCAGAGCTTAGAGATAGCTCTGTGTCAAGAGAAATACATGTAAAGCCATATATTATCT 64370

QY      420 CCAAGTTTTTTAGTAGCCACATTAAACAGTAAACAAAAAGGCTGGGCGCACTGCTCACACC 479
DB      64369 TAAATTTTCTAGTGTCTCTCATTTAAAAAAGTGTAAAGAGCCGGGCGCACTGTCTCACATC 64310

QY      480 TGTATATCCACGACCTTTGGGAGGCTGAGGCGAGGATATCACTTTGTGTAGAGATTGAG 539
DB      64309 TGTATATCCACGACCTTTGGGAGGCGGAGGCGGAGATCAC--GAGTCAAGAGATCAAG 64252

QY      540 ACTAGCCTTGGCCAAATGAGCCGAAACTGTGTCTCTAAACAAAAAATACAAATTAAGCTGG 599
DB      64231 ACCATCTTGGCTTAACATGTGTAAACCCGCTCTTTACTTAAATAATCAAAAATTAAGCCAG 64192

QY      600 CATGTGGCGGGCGGCTGTAACTTCAGCTGTCTCAGAGGCGGAGACACAAGATCACTTA 659
DB      64131 COTGTGGCAAGCACCTGTAGTCCAGCTGCTCGGAGGCTGGGCGAGGATCACTGG 64132

QY      660 AACCCAGAGAGTGTGAGGTTGCACTAGTACGTAGATGTGCACT--CACTCAACCTTGGAG 718
DB      64131 AACCCGAGGAGAGGAGGTTGCACTAGTAAACGAGATGCGGCACTGCACTTACCTTGGAGT 64072

QY      719 ACAGAGTACACTTT--TGTCCTAAACAAAAAAGAAAAAACAAGTAAACAAAGAAAGAGTG 777
DB      64071 ACAGAGAGACTACATCTCAACACAAACAAACAAACAAATTAATAATTAAGAGAG 64012

QY      778 AAGTTAATTAAATNACCAATGTATGCCAAATACATCATTTCAAAGTATTAATAT 837
DB      64011 TAGGTAGAAATGATTTTATNGTTTAT--CTAACCAATATATCCAAATATTTCAATTCA 63954

QY      838 AAAACAATTATGATGAGATGACTTAACTTTCTTTCTGTTTTCATATTAAGCTTTGAA 897
DB      63953 ACATGTATCAATATNAAAAAATTAAGATACCTTATATTTCTTTTTCACACTAAGATTTTGA 63894

QY      898 AGTAGTATATATGTTATGCTGACGACATCTCAATTGGAGCTAGTACATTTCAGGTG 957
DB      63893 AATGTGATATATTTTAACTTATAGCATATCTCAATTCAGAAATGCTCATAGTGTG 63834

QY      958 CTCA 961
DB      63833 GGCA 63830

RESULT 12
US-09-949-016-15779/c
; Sequence 15779, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO010107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15779
; LENGTH: 265038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
;

```



```

Db      107346 CATGTGCGCCACACCTGTATCCAGCTACTCAAGAGTGAGGACACAAAGATCACTTG 107405
Qy      660 AACCCAGAGGTGAGGTTCAGAGAGCTAGATCGGCCACT-CACTCCAACTGGGAG 718
Db      107406 AACCCGGGAAGCGGAGGTTCAGAGAGCCGAGACGAGCCACTGTACTCCAGCCTGTGTG 107465
Qy      719 ACAGAGTACACTTTTGTCTCAAAAAGAAAAAAGAAAAAGTAAAGAAAGAAAGAGTGA 778
Db      107466 ACAGAGCAAGACTGTCTCTCAAAAAAAGAAAAAATCAAGAGAAAGAAACAAATGGA 107525
Qy      779 AGTTAACTTTAATTAACCAATGTAATCCCAATACATCAATTCATTAAGTGAATTAATA 838
Db      107526 ATTATTTTACTAATTTATTTATTTAACCCAGTACATCCAAAGATTAATCAATTCAAAATG 107585
Qy      839 AAACATTAATGATGAGATCTTACATTCCTTTCTGTTCATTTATTAAGTCTTTGAA 898
Db      107586 CAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 107645
Qy      899 GTGAGTATATATGTATGCTGACAGACATCTCAATTTGGAC 940
Db      107646 TCATGTGTGTGTTTACACTTACAGCCCAATTTAATTCGAC 107687

```

RESULT 15

```

US-09-949-016-14170
; Sequence 14170: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14170
; LENGTH: 109159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14170

```

Query Match 11.3%; Score 226.2; DB 4; Length 109159;

Best Local Similarity 68.4%; Pred. No. 2.5e-48; Mismatches 158; Indels 7; Gaps 3;

```

Qy      425 TTTTGTAGTACCAATTAAACAGGTAA-----AAAAGCTGGGCGGAGTGTCTACACC 479
Db      107167 TTTCTAGTGGCCACATTTTAAACATTAAGAAAGAAACAGCCAGCGTGTGCTCATGCC 107226
Qy      480 TGTAAATCCAGCACTTTGGAGGCTGAGGAGGAGCAATCACTTTGCTCAGAGTTGAG 539
Db      107227 TGTAAATCCAGCACTTTGGAGGCTGAGGAGGAGCAATCACTTTGAGCCAGAGTTCAAG 107286
Qy      540 ACTAGGCTGGCCAAATGAGGAAATCTGTCTCTAAAAAAAATACAAAAAATTAAGCTGG 599
Db      107287 ACCAGCTGTGTAACATGTGTGAATCTGTCTCT-ACATAAAATACAAAAATTAAGCCAGG 107345
Qy      600 CATGTGCGGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACAAAGAAATCACTTA 659
Db      107346 CATGTGCGGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACAAAGAAATCACTTG 107405
Qy      660 AACCCAGAGGTGAGGTTCAGAGAGCTAGATCGGCCACT-CACTCCAACTGGGAG 718
Db      107406 AACCCGGGAAGCGGAGGTTCAGAGAGCCGAGACGAGCCACTGTACTCCAGCCTGTGTG 107465

```

```

Qy      719 ACAGAGTACACTTTTGTCTCAAAAAGAAAAAAGAAAAAGTAAAGAAAGAAAGAGTGA 778
Db      107466 ACAGAGCAAGACTGTCTCTCAAAAAAAGAAAAAATCAAGAGAAAGAAACAAATGGA 107525
Qy      779 AGTTAACTTTAATTAACCAATGTAATCCCAATACATCAATTCATTAAGTGAATTAATA 838
Db      107526 ATTATTTTACTAATTTATTTATTTAACCCAGTACATCCAAAGATTAATCAATTCAAAATG 107585
Qy      839 AAACATTAATGATGAGATCTTACATTCCTTTCTGTTCATTTATTAAGTCTTTGAA 898
Db      107586 CAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 107645
Qy      899 GTGAGTATATATGTATGCTGACAGACATCTCAATTTGGAC 940
Db      107646 TCATGTGTGTGTTTACACTTACAGCCCAATTTAATTCGAC 107687

```

Search completed: February 27, 2005, 05:14:33
Job time : 318.935 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:12 ; Search time 1077.07 Seconds
(without alignments)
11008.461 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagagcccccag.....tcgaggggggaaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|---------------------|-------------------|
| 1 | 252.6 | 12.6 | 52216 | US-09-747-810-1 | Sequence 1, Appl1 |
| 2 | 252.4 | 12.6 | 56098 | US-10-087-192-1612 | Sequence 1612, Ap |
| 3 | 240.2 | 12.0 | 55827 | US-09-813-133A-3 | Sequence 3, Appl1 |
| 4 | 240.2 | 12.0 | 55827 | US-10-212-877-3 | Sequence 3, Appl1 |
| 5 | 239 | 11.9 | 1980090 | US-10-719-993-6815 | Sequence 6815, Ap |
| 6 | 239 | 11.9 | 1980090 | US-10-741-600-17676 | Sequence 17676, A |
| 7 | 235.6 | 11.8 | 27189 | US-10-087-192-1828 | Sequence 1828, Ap |
| 8 | 235.6 | 11.8 | 84105 | US-10-741-601-5637 | Sequence 5637, Ap |
| 9 | 229.2 | 11.5 | 653122 | US-10-087-192-226 | Sequence 226, Ap |
| 10 | 227.4 | 11.4 | 23071 | US-09-764-864-1673 | Sequence 1673, Ap |
| 11 | 226.2 | 11.3 | 133955 | US-10-087-192-1984 | Sequence 1984, Ap |

| | | | | | |
|----|-------|------|--------|----------------------|---------------------|
| 12 | 225.4 | 11.3 | 325791 | US-09-768-185A-1 | Sequence 1, Appl1 |
| 13 | 225.2 | 11.3 | 35641 | US-09-962-436-306 | Sequence 306, Ap |
| 14 | 225.2 | 11.3 | 35641 | US-09-880-107-2225 | Sequence 2225, Ap |
| 15 | 224.8 | 11.2 | 629 | US-10-027-632-252422 | Sequence 252422, Ap |
| 16 | 224.8 | 11.2 | 629 | US-10-027-632-252423 | Sequence 252423, Ap |
| 17 | 224.8 | 11.2 | 629 | US-10-027-632-252422 | Sequence 252422, Ap |
| 18 | 224.8 | 11.2 | 629 | US-10-027-632-252423 | Sequence 252423, Ap |
| 19 | 224.8 | 11.2 | 629 | US-10-681-818-155 | Sequence 155, Ap |
| 20 | 224.8 | 11.2 | 629 | US-10-681-818-156 | Sequence 156, Ap |
| 21 | 221.6 | 11.1 | 2591 | US-10-755-889-83 | Sequence 83, Appl |
| 22 | 221.2 | 11.1 | 84409 | US-10-741-601-5696 | Sequence 5696, Ap |
| 23 | 221.2 | 11.1 | 84409 | US-10-741-600-17771 | Sequence 17771, A |
| 24 | 220.8 | 11.0 | 66686 | US-09-736-960-86 | Sequence 86, Appl |
| 25 | 220.8 | 11.0 | 96587 | US-09-997-722-250 | Sequence 250, Ap |
| 26 | 220.6 | 11.0 | 203264 | US-10-087-192-988 | Sequence 988, Ap |
| 27 | 220.2 | 11.0 | 126990 | US-10-717-597-68 | Sequence 68, Appl |
| 28 | 220 | 11.0 | 32146 | US-09-764-860-797 | Sequence 797, Ap |
| 29 | 220 | 11.0 | 32146 | US-10-074-095-797 | Sequence 797, Ap |
| 30 | 220 | 11.0 | 32146 | US-10-212-872-797 | Sequence 797, Ap |
| 31 | 220 | 11.0 | 32248 | US-09-764-860-802 | Sequence 802, Ap |
| 32 | 220 | 11.0 | 32248 | US-10-074-095-802 | Sequence 802, Ap |
| 33 | 220 | 11.0 | 32248 | US-10-212-872-802 | Sequence 802, Ap |
| 34 | 220 | 11.0 | 96589 | US-10-052-482-214 | Sequence 214, Ap |
| 35 | 219.2 | 11.0 | 100267 | US-10-240-425-1470 | Sequence 1470, Ap |
| 36 | 218.6 | 10.9 | 4329 | US-09-764-891-8960 | Sequence 8960, Ap |
| 37 | 218.6 | 10.9 | 43159 | US-10-741-600-17853 | Sequence 17853, A |
| 38 | 218.6 | 10.9 | 16236 | US-10-719-993-6805 | Sequence 6805, Ap |
| 39 | 218.6 | 10.9 | 366803 | US-10-719-993-6805 | Sequence 6805, Ap |
| 40 | 218.4 | 10.9 | 2312 | US-09-764-877-3606 | Sequence 3606, Ap |
| 41 | 218.4 | 10.9 | 2312 | US-10-242-515-3606 | Sequence 3606, Ap |
| 42 | 218 | 10.9 | 2166 | US-10-027-632-98060 | Sequence 98060, A |
| 43 | 218 | 10.9 | 2166 | US-10-027-632-980722 | Sequence 102722, A |
| 44 | 218 | 10.9 | 2166 | US-10-027-632-98060 | Sequence 98060, A |
| 45 | 218 | 10.9 | 2166 | US-10-027-632-102722 | Sequence 102722, A |

ALIGNMENTS

| | | | | | |
|--|-------|--|-------|-------------------|--|
| RESULT 1 | | | | | |
| US-09-747-810-1 | | | | | |
| Sequence 1, Application US/09747810 | | | | | |
| Patent No. US20020012903A1 | | | | | |
| GENERAL INFORMATION: | | | | | |
| APPLICANT: Goldman, Steven A. | | | | | |
| TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGE | | | | | |
| FILE REFERENCE: 19603/3580 | | | | | |
| CURRENT APPLICATION NUMBER: US/09/747, 810 | | | | | |
| CURRENT FILING DATE: 2000-12-22 | | | | | |
| PRIOR APPLICATION NUMBER: 60/173,003 | | | | | |
| PRIOR FILING DATE: 1999-12-23 | | | | | |
| NUMBER OF SEQ ID NOS: 2 | | | | | |
| SEQ ID NO 1 | | | | | |
| LENGTH: 52216 | | | | | |
| TYPE: DNA | | | | | |
| ORGANISM: Homo sapiens | | | | | |
| US-09-747-810-1 | | | | | |
| Query Match | | | | | |
| Beat Local Similarity 12.6%; Score 252.6; DB 9; Length 52216; | | | | | |
| Matches 403; Conservative 0; Mismatches 174; Indels 6; Gaps 4; | | | | | |
| Db | | | | | |
| Qy | 417 | TTTCACTTTTATAGTACCATTAACAGTAAAGCTGGCGAGTGGCTCAC | 476 | Sequence 1, Appl1 | |
| Db | 41515 | TTTAAATTTTTCATGAGCACTTAA--AAGCATAAAGCGCGGCGGTCTCAC | 41572 | Sequence 3, Appl1 | |
| Qy | 477 | ACCTGTATCCAGCATCTTTGGAGGCTGAGCGAGATCCTTTGTCTGAGATT | 536 | Sequence 1828, Ap | |
| Db | 41573 | TCCGTATCCAGCATCTTTGGAGGCTGAGCGAGATGATCCTGAGGTCTGAGATTG | 41632 | Sequence 5637, Ap | |
| Qy | 537 | GAGCTAGCTGGCGAACATGGGAACTGTCTCTTAAAAAAATTAATTAATGACC | 596 | Sequence 226, Ap | |

```
Db 41633 GAGACAGCTGCGCAACAGGGTGAAGCCATGCTCT-ACATAAAATACAAAATTAAGCT 41691
QY 597 TGGCAGTGGGGGGGGCCCTGTATCTCAGCTCTCGAGGCGCGAGACAAGATC 656
Db 41692 GGGTGTGGTGGGGCTGTATCTCAGCTCTCGAGGCGCGAGACAAGATC 41751
QY 657 TTAACCCAGAGGTGAGGTTGTCAGTGAAGTATCGTGCACCT-CACTCCAACTGG 715
Db 41752 TTAAAGAGAGAGGTGAGGTTGTCAGTGAAGTATCGTGCACCTCGAGCTGG 41811
QY 716 GAGACAGAGTGAACCTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAAAGG 775
Db 41812 GTAAAGAGTGAACCTTCATCTCAAAAAAATAAATAATTAATTAATAAATA 41871
QY 776 TGAAG--TTAACTTAAATACCAATGTATCCAAATACATTTCAAGTAAATTA 833
Db 41872 AAAGACATTAATGAAGCAGTGAATTTATTTAATATATATCAAAAAATTAAGCTT 41931
QY 834 ATATAAACCAATTATGATAGATTAATCTTATCTTTCTGTTTCATATTAAGCTT 893
Db 41932 CAACATGATATCAATGTAATTAATATATATCACTGTATTTTCACTTTCTGCTT 41991
QY 894 TGAAGTGAATATATATGTTATGCTGAACAGACATCTCAATTTGGAAGTAACTTGA 953
Db 41992 TGATATCCATGATATTTTGTGACCTTACAGCACTGGTATGTTGGGCGAGCTCATCTCA 42051
QY 954 GGTGCTCAGTACGACATGAGCTAGCAAGTATCTGATTAAGAT 996
Db 42052 AGTGTCTAGTACGACAGCTGGTGAAGTGTCACTTTATTAAGAT 42094
```

RESULT 2

```
US-10-087-192-1612/C
; Sequence 1612, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1612
; LENGTH: 56098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(56098)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1612
```

```
Query Match 12.6%; Score 252.4; DB 13; Length 56098;
Best Local Similarity 70.8%; Pred. No. 2.8e-50;
Matches 408; Conservative 0; Mismatches 156; Indels 12; Gaps 5;

QY 440 TTTAAACAGTAAAGGCTGGCGCAGTGGCTCACACTGTAATCCACGACTTTGGG 499
Db 26680 TTTAAACAGTAAAGGCTGGCGCAGTGGCTCACACTGTAATCCACGAGCTTTGGG 26621
QY 500 AGGCTGAGGAGCAGATCACTTTGGTCAAGATTGAGACTAGAGCTGGCGCAACATGCG 559
Db 26620 AGGCTGAGGAGCAGATCACTTTGGTCAAGATTGAGACTAGAGCTGGCGCAACATGCG 26561
QY 560 GAAACTCTGTCTATAAAAAATATACAAAATTAAGCTGAGTGGCGGCGCTGTGA 619
```

```
Db 26560 GAAACCCATCTCT-ACATAAAATAGAAAAGTTAGCCGATGTGGGCGACAGCTTGTGA 26502
QY 620 ATCTAGCTGCTCAGAGGCGCGAGACAGAAATACATTAACCCAGAGGTGGAGTTGG 679
Db 26501 GTCCAGACGCTCGGAGGCGAGACAGAAATACATTAACCCAGAGGTGGAGTTGG 26442
QY 680 CAGTGAAGTGAATGCTGCCA-CTCACTTCAACTGGAGAGAGAGTGAACATTTGTTCT 738
Db 26441 CAGTGAAGTGAATGAGTATCGCATCTCACTCGGAGAGAGAGAGTGAACATTTGTTCT 26382
QY 739 CAAAAAGAAAAAACAAGTAAAAAAGAAACAGTGAAGTAACTTTAATTAACCAA 798
Db 26381 AAAAAACAAAAAATAA-----CAAAACAAAAACGATATATTTAATTTAAACAATATTA 26328
QY 799 TGAATCCCAATATACATCACTTCAAGTGAATTAATATTAATAAACAATTAATGAATGATA 858
Db 26327 TTTATTTAAACCGTATATCTAAAAATGTCATTTGAATCAATGTAATGAATGAACTTTTGA 26268
QY 859 CTTTACATCTTTCTTCTGTTTCATA-TTAAGCTTTGAAGTGAATATATATGTTATGC 917
Db 26267 ATGGAATATTTTACATCTTTTAAAGATTAAGTCTTTGAAGTGAATGTTATATACAC 26208
QY 918 TGACAGCACA--TTCATTTTGAAGTGAATCACTTCAAGTGTCTCAGTGAAGCAGATGTG 974
Db 26207 TTAATGACACAACTTCTCAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 26148
QY 975 GCTAGCAGTACTGATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1010
Db 26147 GCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 26112
```

RESULT 3

```
US-09-813-133A-3/C
; Sequence 3, Application US/09813133A
; Publication No. US20020137179A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001173
; CURRENT APPLICATION NUMBER: US/09/813,133A
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Human
US-09-813-133A-3
```

```
Query Match 12.0%; Score 240.2; DB 9; Length 55827;
Best Local Similarity 68.6%; Pred. No. 2.7e-47;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTCAGAGTGTGTTTGTAGTAATATGCTGAGCTGTTATGTCATTTCCA 422
Db 46943 TGAATCTCAGACAGGCTTGTATTAAGAAATATATAGCAAGTCAATATTTACTTTAA 46884
QY 423 GTTTTCTAGTACCAATTAATAACAGTAAAA--AAGCTGGGCGCAGTGGCTCACACCT 480
Db 46883 ATTTTCAAGTATCACTTTTAAGAAAAATTAAGAAAGGCGAGGCGAGGCTCACACCT 46824
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGCGAGCAATCACTTTGTCAGAGTTTGAAG 540
Db 46823 ATTAATCCAGCACTTTGGAGGCTGAGGCGAGGCGAGTCACTTGAAGTCAAGAGTTTGAAG 46764
QY 541 CTAGCTGGCGCAACATGGGAAACTCTGTCTTAATAAAAAATTAACAAAATTAAGCTGGC 600
Db 46763 CCAAGGCGCAACATGGGAAACCCGCTCTTACCAAAA--TACAAAATTAAGCCAGGC 46706
QY 601 ATGTGGCGGCGCTGTATCTCAGTGTCTCAGAGGCGGAGACAAAGAAATCACTTAA 660
```

```
Db 46705 ATAGTGTGACGCTGTAAATCCAGCTACTGTGGAGGCTGAGGACGAAATCTTGGA 46646
Qy 661 ACCGAGAGGTGAGGTGGTTCAGTGAAGTGTGCTCACT-CACTCCAACTGGGAGA 719
Db 46645 ACCGAGAGGACGAGGTGGTTCAGTGAAGTGTGCTCACTCACTCCAACTGGGAGA 46586
Qy 720 CAGAGTGACA-CTTTGTCTCAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAGTGA 778
Db 46585 CAGAACAGACCTCTGCCACACACACACACACACACACACACACACACACACACACAC 46526
Qy 779 AGTTAACTTAAATTAACCAATGATCCCAATATACATCAATTCCTCAAGTGAATATATA 838
Db 46525 TTGTAATTTAATTAACCAATATGATACAAAGTATCA----TTTCAAGTGAATATATA 46470
Qy 839 AAACAATTAATGATGATGATCTTCACTCTTTCTGTTTCAATTAATTAAGTCTTTGAA 898
Db 46469 AA--AATTATCTTGAATAATTTAATGATCTT-----TTTCAATTAAGTCTTCAAT 46417
Qy 899 GTGAGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
Db 46416 TCTGCTGTGATTTAATTAACCAATGATGATGATGATGATGATGATGATGATGATGAT 46357
Qy 959 TCAATAGCCACATGTGCTAGCACTTACTGATTTGA 995
Db 46356 TGATAGCCACATGTGCTAGCACTTACTGATTTGA 46320
```

```
RESULT 4
US-10-212-877-3/c
; Sequence 3, Application US/10212877
; Publication No. US20030017574A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1001173D1V
; CURRENT APPLICATION NUMBER: US/10/212,877
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/813,133
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-877-3
```

```
Query Match 12.0%; Score 240.2; DB 14; Length 55827;
Best Local Similarity 68.6%; Pred. No. 2.7e-47;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;
Qy 363 TCATACCTGAGTAGTGTGTTAGTAAATATGCTGAGCTGCTTATGTCATTCCCA 422
Db 46943 TGAATCTCTGACACGAGGTGTGTTAATGAAATATATGCAATGTCATATTTAATTAA 46884
Qy 423 GTTTTATAGTACCATTAATAACAGGTAA--AAGCTGGGCGAGTGGCTCACACT 480
Db 46883 ATTTTCAAGTTACCTTTAAGAAATAAATAAGAGCGGACGCTGGCTCACACT 46824
Qy 481 GTATATCCAGCATTTTGGAGGCTGAGGACGAGATCACTTTGGTCAAGAGTTTGA 540
Db 46823 ATATATCCAGCATTTTGGAGGCTGAGGACGAGATCACTTTGGATCAGAGTTTGA 46764
Qy 541 CTAGCTGGGCAACATGCGGAACTGTCTCTTAAAAAATAAATAAATAAATAAATAAATAA 600
Db 46763 CCAAGGAGGCAACATGCGGAACTGTCTCTTAAAAAATAAATAAATAAATAAATAAATAA 46706
Qy 601 ATGATGAGGAGGCGCTGTAATCTTCAAGTGTCTCAGAGGCGGAGACACAAGATCACTTAA 660
Db 46705 ATGATGAGGAGGCGCTGTAATCTTCAAGTGTCTCAGAGGCGGAGACACAAGATCACTTAA 46646
```

```
Qy 661 ACCGAGAGGTGAGGTGGTTCAGTGAAGTGTGCTCACT-CACTCCAACTGGGAGA 719
Db 46645 ACCGAGAGGACGAGGTGGTTCAGTGAAGTGTGCTCACTCACTCCAACTGGGAGA 46586
Qy 720 CAGAGTGACA-CTTTGTCTCAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAGTGA 778
Db 46585 CAGAACAGACCTCTGCCACACACACACACACACACACACACACACACACACACACAC 46526
Qy 779 AGTTAACTTAAATTAACCAATGATCCCAATATACATCAATTCCTCAAGTGAATATATA 838
Db 46525 TTGTAATTTAATTAACCAATATGATACAAAGTATCA----TTTCAAGTGAATATATA 46470
Qy 839 AAACAATTAATGATGATGATCTTCACTCTTTCTGTTTCAATTAATTAAGTCTTTGAA 898
Db 46469 AA--AATTATCTTGAATAATTTAATGATCTT-----TTTCAATTAAGTCTTCAAT 46417
Qy 899 GTGAGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
Db 46416 TCTGCTGTGATTTAATTAACCAATGATGATGATGATGATGATGATGATGATGATGAT 46357
Qy 959 TCAATAGCCACATGTGCTAGCACTTACTGATTTGA 995
Db 46356 TGATAGCCACATGTGCTAGCACTTACTGATTTGA 46320
```

```
RESULT 5
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6815
```

```
Query Match 11.9%; Score 239; DB 18; Length 1980090;
Best Local Similarity 72.6%; Pred. No. 2.9e-46;
Matches 416; Conservative 0; Mismatches 125; Indels 32; Gaps 7;
Qy 437 ACATTAACAAGGTAATAAAGGCTGGGCGGAGGCTGACACCTGTATCCAGACCTTT 496
Db 539204 AATTAAGAAAGAAATAGAGAGGCTGGGCGGAGGCTGATCTGTATCTCCAGACCTT 539145
Qy 497 GGGAGGCTGAGGAGGAGGAGTCACTTTGGTCAAGAGTTTGAATAGCTTGGCCAAAT 556
Db 539144 GGGAGGCTGAGGAGGAGGAGTCACTTTGGTCAAGAGTTTGAATAGCTTGGCCAAAT 539085
Qy 557 GGGAAATCTGTCTCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 616
Db 539084 GGTGAAATCCCGTCTCT-CTTAAAAATGACAAATAAATAGCCGCGGTGGGAGGTGCT 539026
Qy 617 GTATCTCAGCTCTCAGAGGCGGAGACACAAGATCACTTAACTCCAGAGGTGGAGG 676
Db 539025 GTATCTCAGCTCTCAGAGGCGGAGAGG-CGAGGAGAAATTTGCTCAATCCAGAGGTGAGG 538967
Qy 677 TTGCACTGAGCTGAGATCTGCTCACT-CACTCCAACTGGGAGACAGAGTGAATTTTG 735
Db 538966 TTGCACTGAGCTGAGATTTGCTCACTGCACTCCAGCTGGGAAACAGAGAGATCTCACT 538907
```

| | | | |
|----|--------|--|--------|
| Oy | 736 | ICTCAAAAGAAAAAAAACAAGTAAAAAGAGGTAACTTAATTAAAC | 795 |
| Db | 538906 | TGAAAAAAGAAACAGAGAAATTAGCAATA-----TATTTCAATTACT | 538859 |
| Oy | 796 | CAAGTATCCCAATACATCTTCAAGGTAAATTAATATAAACAATTATGAATGAG | 855 |
| Db | 538858 | CAATGATCCAAATTATGGTCATTGACATGTAATCAATATATGAG-----TTTATATGAG | 538804 |
| Oy | 856 | ATACTTCATCTTTCTGTGTTTTCATTTAAGTCTTGAAGAGATATAAGTTAT | 915 |
| Db | 538803 | ATAGTCTATGTTCTTTTTCCTAAG-----TCTGAATTTGGGTGTAT--TTAC | 538756 |
| Oy | 916 | GCTGACAGACATCTCAATTTTGGACTAGCTACATTTCAAGTGCTCAGTACCCACATGTGG | 975 |
| Db | 538755 | ACTGACAGACATCTCAATTTTAGACTAGCCACATTTCCAGTGTGATGGCTGCATGTAC | 538696 |
| Oy | 976 | CTAGCAGTTACTGTATTTGATGGCAGGATCTTA | 1008 |
| Db | 538695 | CTGGTGGCTACCACTGGACAGCAACAGGTCTTA | 538663 |

```

RESULT 6
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169a1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-741-600-17676

```

| Query Match | 11.9% | Score 239 | DB 19 | Length 19809090 |
|-----------------------|----------------|--|-----------|-----------------|
| Best Local Similarity | 72.6% | Pred. 2.9e-46 | | |
| Matches 416 | Conservative 0 | Mismatches 125 | Indels 32 | Gaps 7 |
| Qy | 437 | ACATTAAACAGGTAAAAAGGCTGGGCGCAGGGCTCAACCTGTAAATCCAGCACTT | 496 | |
| Db | 539204 | AAATGAAAGAAATAGAGAGGCTGGGCGTGGCTCAAGCTGTAAATCCAGCACTT | 539145 | |
| Qy | 497 | GGGAGGCTGAGGCGAGGAGATCACTTTGGTCAGAGATTGAGATGACCTGGCCAAAT | 556 | |
| Db | 539114 | GGGAGGCCAAGGCGGGCAGATCACTGAGGTGAGAGTTTGAGACCAAGCTGGCCAAAT | 539085 | |
| Qy | 557 | GGCGAACTCTGTCTCTAAAAAAAATACAAAAATTAGCGCTGATGAGTGGCGGCGCCT | 616 | |
| Db | 539084 | GGTGAACCCCGTCTCT-ACHTAAAGTACAAAAATTAGCCGGCGGTGGCGAGGTGCT | 539026 | |
| Qy | 617 | GTAATCTCAGCTGCTCAGGAGGCGAGACACAAGATCACTTAAACCGAGAGGTGAGG | 676 | |
| Db | 539025 | GTAATCTCCAGCTACTCAGGAGG-CGAGCGAGAAGATTGCTCAAAACCCAGAGGTGAGG | 538967 | |
| Qy | 677 | TTGCAGTGAGCTGAGATCGTGCACCT-CACTCCAACTTGGGAGACAGAGTGACACTTTTG | 735 | |
| Db | 538966 | TTGCAGTGAGCCAAAGTTTGCCACTGCACTCCAGCTGGGAAAAGAGGAGAGCTCAGT | 538907 | |
| Qy | 736 | TCTCAAAAAGAAAAAAAACAAGTAAAAAAGAAACAGGTGAAGTTAACTTTAATACC | 795 | |
| Db | 538906 | TGAAAAAAGAAAGAAACAGAGAAATTAGCAATA-----TATTTCAATTACAC | 538859 | |
| Qy | 796 | CAATGTATCCAAATACATCATTTTCAAGGTAAATTAATATAAACAATTTATGAATGAG | 855 | |

| Accession | Sequence | Position |
|-----------|--|----------|
| D6 | 538858 CAAATGATCCAAATATATGTCATTTTACCATGTAATCAATATAGA-----TTTAAATGAG | 538804 |
| Oy | 856 ATACTTACATCTCTTTCTTGTTTCATATTTAGCTTTTGAAGTAGATATATGTTAT | 915 |
| D6 | 538803 ATAGTCATGTTCTTTTCTTCTAAG-----TCTGAATTCGGTGTGAT-TTAC | 538756 |
| Oy | 916 GCTGACAGCAATCTCAATTTGSACTAGCTACATTTCAAGTGTCTAGTACCAATGTGG | 975 |
| D6 | 538755 ACTGACAGCAATCTCAATTTTGACTAGCCACATTTCCAGTGTCTAGTGTGCTGCAATGTAC | 538656 |
| Oy | 976 CTACAGTTACTGTAATTTGATGAGCAGGATCTTA | 1008 |
| D6 | 538695 CTGGTGGCTACCAACTGACAGCAGCAGAGTCTTA | 538663 |

```

RESULT 7
US-10-087-192-1828/c
; Sequence 1828, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1828
; LENGTH: 27189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1828

```

| Query Match | 11.8%; | Score 235.6; | DB 13; | Length 27189; |
|-----------------------|-----------------|---|------------|---------------|
| Best Local Similarity | 68.8%; | Pred. No.2, 5e-46; | | |
| Matches 385; | Conservative 0; | Mismatches 164; | Indels 11; | Gaps 4; |
| Qy | 439 | ATTAAACAGTATAAAGGCTGGGCGCAGTGGCTCACACTGTAATCCAGCACTTTGG | 498 | |
| Db | 11830 | AGTAAATATGCTAAATATGGGCGCGGCGATGTAATCTATCTTGTATCCACACTATGA | 117711 | |
| Qy | 499 | GAGGCTGAGGCGAGGAGTACCTTTGTGCAGAGATTGAGACTGAGCTGGCCGCAATGG | 558 | |
| Db | 11770 | GAGGCGAGGTTGGGCAAGATCACTGAGGTCAGAGATTGAGACAGGCTGGCCCAATGG | 117111 | |
| Qy | 559 | CGAACTCTGTCTCTATAAAAAAATACAAAAATTAGCCTGGCATGTTGGCGGCGCTGT | 618 | |
| Db | 11710 | TGAAACCCGCTCT-ACATAAGATATAAAATTAGCTGGAGTGTGGTGGCTCTGT | 116522 | |
| Qy | 619 | AATCTACGCTGCTCAGAGCGCGAGACAAATATCACTTAACCCAGAGGTGAGGTT | 678 | |
| Db | 11651 | AATCTCAGCTACTCAAGAGGCTGAGGCAAGAAATTTCTTGAACCCAGAGGCAAGGTT | 115952 | |
| Qy | 679 | GCAGTGAAGTGAATGATGTGCACTCACTCCAACCTGGGAGACAGATGACACTTTTGTCT | 738 | |
| Db | 11591 | GCAGTGAAGCGAGACGTGCACCAATTTGACCTTCAGCATGAATGGCAAGGCGAAAGTCTGT | 115322 | |
| Qy | 739 | CAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTAAAGTTAATCTTAATAACCGA | 798 | |
| Db | 11531 | CAAAAAAAGAAAAA-----AAAAAGCAACAGCTTTTGGCAGATGATCAACATTAC | 114777 | |
| Qy | 799 | TGATCCCAAAATCATCTTTCAAGATGATTAATATATAAACAATTGATGAGATA | 858 | |
| Db | 11476 | TACACTCCCTTCCCTTCACTTTTCAAAAAAATTTTCATGCACTTTAACAAATATTATA | 114177 | |
| Qy | 859 | CTTACATCTCTTTCTGTGTTTCATATTAAGTCTTTGAAAGTGAAGTATATATGTTATGC- | 917 | |

Db 11416 GATTACATCTC-----TTCTTCTCATGTAACTCTTTAAATCTGGGGTAAATTTTAACT 11361

QY 918 TGACAGCACAATCTCAATTTGGACTAGCTACATTTCAAGTGTCTCAGTAGACCAATGTGGCT 977

Db 11360 TAAAGCATGTCTTAATTTGGACTAGACACATTTCAAGTGTCTCAGTAGACAGTGAGAT 11301

QY 978 AGCAGTACTGTATTGGATG 997

Db 11300 AGTGTCACTGTATTAGATG 11281

RESULT 8

```

US-10-741-601-5637/c
; Sequence 5637, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michale et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5637
; LENGTH: 84105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5637

```

| | | | | |
|---------------------------|-------|-------------------|----------|--------------|
| Query Match | 11.8% | Score 235.6 | DB 18 | Length 84105 |
| Best Local Similarity | 67.4% | Pred. No. 4.4e-46 | | |
| Matches 407, Conservative | 0 | Mismatches 189 | Indels 8 | Gaps 5 |

| | | | |
|----|-------|--|-------|
| Oy | 362 | CTCATACCCTAGAGATAGTGGGCTTTAGTAAAGT--AAATGCTAGGCTGTAATGTCATTT | 419 |
| Db | 68837 | CTCAGAGCTTAGAGTAAGCTCTGCTGCAGAGAGAAATACATATGAAGCCACTATATTTATCT | 68778 |
| Oy | 420 | CCAGTTTTTTTAGTAGCCACATTTAAACAGGTAAAAAAAGCTGGGCGCAGTGGCTCACCC | 479 |
| Db | 68777 | TAAATTTTCTAGTTGGCTCTCATTTAAAAAAGTAAAGAGACGGGGCCAGTGTCTCACATC | 68718 |
| Oy | 480 | TGTAATCCCGAGCAATTTGGGAGCGTGAGCGACATCACTTTGGTCAGAGTTTGAG | 539 |
| Db | 68717 | TGTAATCCCGAGCACTTTGGGAGCGCGGTGTGAGATCAC--GAAGTCAAGGAGATCAAG | 68660 |
| Oy | 540 | ACTAGCTGGCGCAATGCGGAAACTGTGTCTCTAAAAAAAAATACAAAAATTAGCGTGG | 599 |
| Db | 68659 | ACCATCTGGGTAAATGATGTAAGAACCCCGCTTTTACTAAATAATCAAAAAATTAGCCAG | 68600 |
| Oy | 600 | CATGTGTGGCGGGCGCCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCATTTA | 659 |
| Db | 68599 | CGTGTGTGGCAGGCACTGTAGTCCCAAGCTCTCGGGGAGCGTGGGAGGAGAGATCACTGG | 68540 |
| Oy | 660 | AACCAGAGAGGTGAGGTTGCAGAGAGCTAGATCTGGCCACT--CACTCCAACTGGGAG | 718 |
| Db | 68539 | AACCCGGAGGCAAGGTTGAGTGAAGTGAACCGAGATCGGCACTGCACCTCAGCTGGGGT | 68480 |
| Oy | 719 | ACAGAGTGACATTT-TGTCTCAAAAAGAAAAAAMAAAACAGTAAMAAAAGAAACAGGTG | 777 |
| Db | 68479 | ACAGAGAGAGCTCATCTCTCAACCAACAAACAAACAAACAAATATATATATTAAGAGA | 68420 |
| Oy | 778 | AAGTTAATCTTAATAAACCAATGTATCCCAATATCAATCATCTTCAAGTGTATTAATAT | 837 |
| Db | 68419 | TAGGTAGAAATGATTTATGTATTTAT--CTAACCCAAATATATCCAAATATTTCAATTTA | 68362 |
| Oy | 838 | AAAACAATTATGATGATGATACATTATCATCTTTCTGTCTCATATTAAGCTTTGAA | 897 |
| Db | 68361 | ACATGTATCAATATATAAAATTAAGATACATTATATCTTTTTCACACTAAGATTTGA | 68302 |
| Oy | 898 | AGTGAATATATATGTTATGCTGACAGCAATCTCAATTGGACTAGCTACATTTCAAGGTG | 957 |

| | | | |
|----|-------|---|-------|
| Db | 68301 | AATTGGATATATTTTCACTTATAGCATTCATTCAGATACCTCATGGTGGTG | 68242 |
| Qy | 958 | CTCA | 961 |
| | | | |
| Db | 68241 | GGCA | 68238 |

RESULT 9

```

US-10-087-192-226/c
; Sequence 226, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226
; LENGTH: 653122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (653122)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-226

```

| | | | | |
|-----------------------|-----------------|-------------------|------------|----------------|
| Query Match | 11.5% | Score 229.2; | DB 13; | Length 653122; |
| Best Local Similarity | 69.7% | Pred No. 4.3e-44; | | |
| Matches 404; | Conservative 0; | Mismatches 153; | Indels 23; | Gaps 6; |

| | | | |
|----|--------|--|--------|
| Oy | 440 | TTTAAACAGGTAAAAAAGCTGGGGCAGTGGCTCACACCTGTAAATCCAGCACTTTGGG | 499 |
| Db | 360810 | TTTTAAAAATACATTTGGGCCAGGTGGCTCATGTCTGTAAATCCAGCACTTTGAA | 360751 |
| Oy | 500 | AGCTGAGGCAAGCAGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCAACTGGC | 559 |
| Db | 360750 | AGCGGAGGGCCGGGCAATCAC--GAGGTGAGGAGTTTGAGACCAAGCTCAGCAACTGGT | 360693 |
| Oy | 560 | GAAACTCTGTCTCTAAAAAAAATACAAAATTAGCTGGCATGTGGCGGGCGCTGTAA | 619 |
| Db | 360692 | GAAACCCCGCTCT-ACTAAAAATTCAAAAATTACCGAGCGCTGGTGTGGCACGCTGTAA | 360634 |
| Oy | 620 | ATTCAGTGTGTCAGGAGGCGGAGACAAAGATACCTTAAACCCAGAGGTGGAGAGTTG | 679 |
| Db | 360633 | ATCCAGCACTTGGGAGGCTGGAGCAGAGAGATACCTTGAATCTGGGAGGCGAGAGTTG | 360574 |
| Oy | 680 | CAGTGAGCTGAGATGGTGCCACT-CACTCCAACCTGGGAGACAGAGTGAACATTTGTCT | 738 |
| Db | 360573 | CAATGAACTGAGATTATGCCACTGCACTTCAGCTGGGGGAGCAGAGCAAGACTCCATCTG | 360514 |
| Oy | 739 | CAAAAAGAAAAAAAACAAGTAAAAAGAAACAGTGAAGTAACTTTAAATACCCA | 798 |
| Db | 360513 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAATTGG-----ATTTAATTTAA | 360466 |
| Oy | 799 | TGTATCCCAATPACAAATCATTTCCAAAGTGTAATTAATAAACAATTGAATGAGATA | 858 |
| Db | 360465 | TATATCCAAAAGTTTAAATTTCAACATGCAACTATATATAAAGATGAATGATGTA | 360406 |
| Oy | 859 | CTTTACATCTCTTTCTGTTTTCATATTAAGCTTTGAAGAGATATATATATGTCT | 918 |
| Db | 360405 | ACCTTCTTCTCTTTTTC---TACTAAGCTTTGATGTCAGACATATATTTTATACA | 360349 |
| Oy | 919 | GACAGCACTCTCAATTTGACTAGCT---ACATTTGAGGTGCTCAGTACGCAACTGTG | 974 |

Db 360348 TATGCCCTATTGATTGGGGCTAGCTAGCCAGCTTCCAGGGCTCAAGACCATCTG 360289
QY 975 GCTAGCAGTACTGTTATGGATGCA CGGATCTAGAGGGA 1014
Db 360288 GCTGTAGCTACCATAGCAGACAGTGCAGACTTAAGACA 360249

RESULT 10

US-09-764-1673/C
; Sequence 1673, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1673
; LENGTH: 23071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-1673

Query Match 11.4%; Score 227.4; DB 9; Length 23071;
Best Local Similarity 70.1%; Pred. No. 2.4e-44;

Matches 445; Conservative 0; Mismatches 156; Indels 34; Gaps 9;

QY 381 GTGTTTAGTAAATATATGCTGAGCTGCTTATGCTATTCATTTCCAGTTTATAGTACCAT 440
Db 10748 GTGCTAAGGAATATATATGAGCCACATATATTTAAATTTCTGGCTCCACATT 10689
QY 441 TAAAAAG-CTAAAAAGCTGGGCGAGTGGCTCACTGTAATCCAGCATTTGGG 499
Db 10688 AACAAAAAGTAAAAAGAGGCTGGGCGAGTGGCTCACTGTAATCCAGCATTTGGG 10629
QY 500 AGGCTAGGAGCAGATACCTTTGGTCAAGGTTGAGTACCTGGCCAAACATGGC 559
Db 10628 AGGCTAGAGAGCGAGTACCTGAGGCTGGGAGTTCGAGCTTAACCAACATGGA 10569
QY 560 GAAATCTGTCT---CTAAAAAAAATATCAAAAATTTAGCTGAGTGGCGGGCGCT 616
Db 10568 GAAATCCATCTCTACTATAAAAAAAATTTCAAAATTTAGCTGGCGTGGTGGCACAATGCT 10509
QY 617 GTATCTCAGCTGCTCAGAGGCGGAGACACAAATTAATTAATCCAGAGTGGAGG 676
Db 10508 GTATCTCAGCTCAGAGGCGGAGTGGAGGAGAAATCGCTTAACCCGGGAGCGAGG 10449
QY 677 TTGACGAGCTGAGATCGTGCACCT-CACTCCAGCTGGGAGACAGATGACATTTTG 735
Db 10448 TTGCGGTAGCGAGATTCGCGCATTTGCACTCCAGCTGGGCAACAGCGAACTCCA 10389
QY 736 TCTCAAAAAAGAAAAAAAACAGTAA-----AAAAGAAACAGGTGAAGTTAACTTT 788
Db 10388 TCTCAAAAAAGAAAAAAAACAGTAA-----AAAAGAAACAGGTGAAGTTAACTTT 10329
QY 789 AA-----TAACCAATGATCCCAATTAATCAATTTCAAGTGAATTAATTAATAAC 843
Db 10328 AATAGTTTAAACAGGTATCCATGATATATTTCAACATGATTAACATTA--A 10271
QY 844 AATAGTAAAGATTAATCAATTTCTTTTCTTTTCAATTAAGTCTTTGAAGTGG 903
Db 10270 ACTATTAAGTAACTTTTCAATTTGTT-----TATCAATTTCAAAATCCAT 10223
QY 904 TATATATGTTATGCTG-ACAGCAGATCTCAATTTGAGTACATTTCAAGGTCTCAG 962
Db 10222 TGTGTATTTTCTATCTTCAAGCAGCAACTCAATTTGGG--CAGCAGATTTCAATTTCTCA 10165
QY 963 TAGCCAGATGTGGCTGACGATTAATCTGATTTGATG 997
Db 10164 TAGCCAGAAATGCTATTTGGCTATTTTATTTGTTG 10130

RESULT 11

US-10-087-192-1984
; Sequence 1984, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1984
; LENGTH: 133955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1984

Query Match 11.3%; Score 226.2; DB 13; Length 133955;
Best Local Similarity 68.4%; Pred. No. 1.1e-43;

Matches 357; Conservative 0; Mismatches 158; Indels 7; Gaps 3;

QY 425 TTTTATAGTACCAATTAATAACAGTAA-----AAAAGCTGGGCGAGTGGCTCACACC 479
Db 115167 TTTTATAGTACCAATTTTAAACATTAAGAAAGAAACAGGCGGTGGTCTCATGCC 115226
QY 480 TGTATCCAGCACTTTGGAGGCTGAGGCGAGCAGATCACTTTGGTGAAGATTGAG 539
Db 115227 TGTATCCAGCACTTTGGAGGCTGAGGCGAGCAGATCACTTTGGTGAAGATTGAG 115286
QY 540 ACTAGCTGGGCGCAATGCGGAACCTGTCTCTAAAAAAAATCAAAATTTAGCGCTGG 599
Db 115287 ACCAGCTGGGTAAACATGCTGTAACCTGTCTCT-ACATAAAATCAAAAATTTAGCGCAG 115345
QY 600 CATGTGGGCGGCGCTGTATCTCAGTGTCTCAGAGGCGGAGACACAAGATCACTTA 659
Db 115346 CATGTGGGCGGCACTGTATCCAGTACTCAAGTATGATGAGGACAAAGATCACTTG 115405
QY 660 AACCCAGAGGTGAGGTTCAGTGAAGTGAAGTGTGCACT-CACTCAACTGGGAG 718
Db 115406 AACCCGAGAGGCGAGGTTCAGTGAAGTGAAGTGAAGTGTGCACTCACTGTGTG 115465
QY 719 ACAGAGTGAACATTTTGTCTCAAAAAAGAAAAAAAACAATTAATAAGAAAGAGTGA 778
Db 115466 ACAGAGTGAACATTTTGTCTCAAAAAAGAAAAAAAACAATTAATAAGAAAGAGTGA 115525
QY 779 AGTTACTTTAATAACCAATGATCCAAATATACATCAATTTCAAGTGAATTAATTA 838
Db 115526 ATTATTTTACCAATTTTATTTAACCCAGATATCCAAAGATTTATCAATTTCAAAATG 115585
QY 839 AAACATTAATGAATGATGATCTTAACATTTCTTTTCTTTTCAATTAATTAATTAAT 898
Db 115586 CAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 115645
QY 899 GTGAGTATATGTTATGCTGACAGCAGATCTCAATTTGGAC 940
Db 115646 TCAATGATGTTTATTAACCTTAACGCCCAATTTTAATTTGGAC 115687

RESULT 12

US-09-768-185A-1
; Sequence 1, Application US/09768185A
; Publication No. US20040185439A1
; GENERAL INFORMATION:
; APPLICANT: Casseel, Michael et al

TITLE OF INVENTION: Estrogen receptor beta variants and
TITLE OF INVENTION: methods of detection thereof
FILE REFERENCE: C0000280
CURRENT APPLICATION NUMBER: US/09/768,185A
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 09768185
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 325791
TYPE: DNA
ORGANISM: HUMAN
US-09-768-185A-1

Query Match 11.3%; Score 225.4; DB 11; Length 325791;
Best Local Similarity 82.6%; Pred. No. 2.6e-43;
Matches 270; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 441 TAAAGAGGTAAAGAGGTGGGCGAGTGGCTACACCTGTATCCAGACCTTTGGGA 500
Db 282763 TGAAGAGAAACAGAGAGGCGGCGGCGAGTGGCTACACCTGTATCCAGACCTTTGGGA 282822
Qy 501 GGGTGGAGGAGGAGATCACTTTGGTTCAGAGTTTGAAGTCTAGCTGGCCAGATGCGG 560
Db 282823 GGGCAAGGAGGAGTGGATCATCTGAGGTGAGAGTTCAAGACTAGCTGGCCAGATGCGG 282882
Qy 561 AAATCTGTCTCTTAAATAAATAATTAAGCTGGCATGTGGCGGCGGCTGTAA 620
Db 282883 AAATCTGTCTCTCTTAAATAAATAATTAAGCTGGCATGTGGCGGCGGCTGTAA 282942
Qy 621 TCTCAGCTGCTCAGAGGCGGAGACAGAGATCACTTAACCCAGAGGTGGAGGTTC 680
Db 282943 TCCAGCTGCTCGGAGGCTGAGAGAGAGATGCTTGAACCCAGAGGTGGAGGTTC 283002
Qy 681 AGTAGGTGAGATGCTGCTCACTTCCAACTGGGAGAGAGATGACATTTGTCTC 739
Db 283003 AGTAGGTGAGATGCTGCTCACTTCCAACTGGGAGAGATGACATTTGTCTCA 283062
Qy 740 AAAAAAAAAAAAAAAAACACTAATAA 766
Db 283063 AAAAAAAAAAAAAAAAAAGAGAAA 283089

RESULT 13

US-09-962-436-306
Sequence 306, Application US/09962436
Patent No. US20020081301A1
GENERAL INFORMATION:
APPLICANT: Soppet, Daniel
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-75
CURRENT APPLICATION NUMBER: US/09/962,436
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,082
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/234,924
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 568
SOFTWARE: PatentIn version 3.0
SEQ ID NO 306
LENGTH: 35641
TYPE: DNA
ORGANISM: Homo sapiens
US-09-962-436-306

Query Match 11.3%; Score 225.2; DB 9; Length 35641;
Best Local Similarity 84.4%; Pred. No. 1e-43;
Matches 265; Conservative 0; Mismatches 48; Indels 1; Gaps 1;
Qy 455 AAGCTGGGCGGAGTGGCTCAACCTGTATCCAGACCTTTGGAGGCTGAGGAGGCA 514
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTGAAGTCTGCTCAAAAAAAAAAAAAAAAA 14988

Db 14689 AAGCTGGGCTCAGTGGCTCAGCTGTATATCCAGACCTTTGGAGGCTGAGGAGGCC 14748
Qy 515 GATCACTTTGGTCAGAGAGTTTGAAGTACTACCTGGCCAAATAGCCAAATCTGTCTCTA 574
Db 14749 GATCACTGGAGTCAGAGAGTTTGAAGTACTACCTGGCCAAATAGTGAAGCCCGTCTCTA 14808
Qy 575 AAAAAAAAAATCAAAAATTTAGCTGGCATGTGGCGGCGGCTGTATCTCAGCTGTCTAG 634
Db 14809 TTTAAATAATACCAAAAATTTAGCTGGCATGTGGCGGCGGCTGTATCTCAGCTGTCTAG 14868
Qy 635 GAGGCCGAGACACAGAAATCACTTAAACCCAGAGGTGAGGTTGCAAGTCAAGTCAATC 694
Db 14869 GAGGCTGAGGACAGAGAAATCACTTAAACCCAGAGGTGAGGTTGCAAGTCAAGTCAATC 14928
Qy 695 GTGCACT -CACTTCAACTGGGAGAGAGTGAACATTTTGTCTCAAAAAAAAAAAAA 753
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTCTGCTCTCAAAAAAAAAAAAA 14988
Qy 754 AAACAAGTAAAAA 767
Db 14989 AAAAAAAAACAAAAA 15002

RESULT 14

US-09-880-107-2225
Sequence 2225, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2225
LENGTH: 35641
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L76568
US-09-880-107-2225

Query Match 11.3%; Score 225.2; DB 9; Length 35641;
Best Local Similarity 84.4%; Pred. No. 1e-43;
Matches 265; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Qy 455 AAGCTGGGCGGAGTGGCTCAGACCTGTATATCCAGACCTTTGGAGGCTGAGGAGGCA 514
Db 14689 AAGCTGGGCTCAGTGGCTCAGCTGTATATCCAGACCTTTGGAGGCTGAGGAGGCA 14748
Qy 515 GATCACTTTGGTCAGAGAGTTTGAAGTACTACCTGGCCAAATAGCCAAATCTGTCTCTA 574
Db 14749 GATCACTGGAGTCAGAGAGTTTGAAGTACTACCTGGCCAAATAGTGAAGCCCGTCTCTA 14808
Qy 575 AAAAAAAAAATCAAAAATTTAGCTGGCATGTGGCGGCGGCTGTATCTCAGCTGTCTAG 634
Db 14809 TTTAAATAATACCAAAAATTTAGCTGGCATGTGGCGGCGGCTGTATCTCAGCTGTCTAG 14868
Qy 635 GAGGCCGAGACACAGAAATCACTTAAACCCAGAGGTGAGGTTGCAAGTCAAGTCAATC 694
Db 14869 GAGGCTGAGGACAGAGAAATCACTTAAACCCAGAGGTGAGGTTGCAAGTCAAGTCAATC 14928
Qy 695 GTGCACT -CACTTCAACTGGGAGAGAGTGAACATTTTGTCTCAAAAAAAAAAAAA 753
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTGAAGTCTGCTCAAAAAAAAAAAAA 14988

OY 754 AACACGTAAAAA 767
Db 14989 AAAAAACAAAAA 15002

Db 597 AAAAAAGAAACAGGTGAATTAATTTAAT 628

Search completed: February 27, 2005, 04:52:27
Job time : 1097.07 secs

RESULT 15

US-10-027-632-252422
; Sequence 252422, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252422
; LENGTH: 629
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-252422

Query Match : 11.2%; Score 224.8; DB 13; Length 629;

Best Local Similarity 72.8%; Pred. No. 1.8e-44;
Matches 329; Conservative 1; Mismatches 118; Indels 4; Gaps 3;

OY 340 CAGGACCCCTGTTCTCAGAGCTCAATCCCTAGAGTAGTGTGTTAGTAGAAATPATG 399
Db 181 CAGTACGTTTACATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
OY 400 CTGAGCTGCTTAATGTCATTTCCAGTTTCTTGTAGCCATTAACAGGTAAAAAGGC 459
Db 241 TACAGGCCACACATGCTATTTTAATTTTCTGTAGCCATTAATTAATTAATTAATTAAT 300
OY 460 TGGGCGGAGTGGCTCACACTGTAAATCCAGCACTTTGGAGGCTGAGGCAAGCATCA 519
Db 301 TGGGTGAGTGGCTCAAGCTGTAATCCAGAACTTTGGAGGCAAGGCGGATCA 360
OY 520 CTTTGGTCAGAGTTTGAAGTCTGAGCTGAGCAATGCGAAATCTGCTCTAAAAAA 579
Db 361 --TGAGGTCAAGATGAGATCACTCTGCGCAATATGTAATCCCGTCTCT-ACTAA 417
OY 580 AAAATACAAAAATTAAGCTGAGCATGTGGCGGCGCTGTAACTCAAGCTGCTCAGAGGC 639
Db 418 AAATACAAAAATTAAGCTGAGCATGTGGCGGCGCTGTAACTCAAGCTGCTCAGAGGC 477
OY 640 CGAGACACAAAGATCACTTAAACCCAGAGAGTGAAGTTGAGTGAATCGTGCC 699
Db 478 TGAGGCGAAGAAATCGCTTAAACCCAGAGAGTGAAGTTGAGTGAATCGTGCC 537
OY 700 ACTCACTCGAAGCTGGAGACAGAGTGAACCTTTTGTCTCAAAAAAGAAAAAAACAA 759
Db 538 ACTGCACTCGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
OY 760 GTAAAAAAGAACAGGTGAAGTTAATTTAAT 791

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 20:26:59 ; Search time 6271.68 Seconds
(without alignments)
12144.524 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagaagccccaag.....tcgcaggggggaaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapexc 1.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68473088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa81: *
9: gb_gsa82: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 227 | 11.3 | 908 | 6 | CD245375 |
| 2 | 223 | 11.1 | 694 | 8 | BZ603010 WHAAX35TF |
| 3 | 221.6 | 11.1 | 1750 | 3 | AF119908 Homo sapi |
| 4 | 220 | 11.0 | 769 | 6 | CD370173 UI-H-FP1- |
| 5 | 219.6 | 11.0 | 500 | 5 | BX491983 DKFZP7810 |
| 6 | 218.4 | 10.9 | 837 | 2 | BE790769 601581950 |
| 7 | 216 | 10.8 | 649 | 8 | B71494 RPTC11-889. |
| 8 | 216 | 10.8 | 1019 | 6 | CD518470 AGENCOURT |
| 9 | 215.6 | 10.8 | 2449 | 3 | CR749224 Homo sapi |
| 10 | 215.4 | 10.8 | 573 | 5 | BX477553 DKFZP686H |
| 11 | 215.4 | 10.8 | 823 | 6 | CD101954 AGENCOURT |
| 12 | 215 | 10.7 | 397 | 1 | AV759632 |
| 13 | 215 | 10.7 | 387 | 5 | BX644875 DKFZP781D |
| 14 | 215 | 10.7 | 633 | 7 | CK905714 1j76h09.x |
| 15 | 214 | 10.7 | 491 | 2 | AM516097 xtc6f05.x |
| 16 | 214 | 10.7 | 4593 | 3 | HM803561 |
| 17 | 213.4 | 10.7 | 517 | 8 | AM264959 CITR1-B1- |
| 18 | 213 | 10.6 | 569 | 5 | AQ490878 RPTC11-2 |
| 19 | 212.8 | 10.6 | 700 | 5 | BX508919 DKFZP686G |
| 20 | 212.6 | 10.6 | 790 | 6 | CD643827 AGENCOURT |
| 21 | 212.4 | 10.6 | 619 | 5 | BX485408 DKFZP686B |
| 22 | 211.8 | 10.6 | 576 | 6 | CD652305 AGENCOURT |
| 23 | 211.8 | 10.6 | 731 | 7 | CN307840 170004182 |
| 24 | 211.8 | 10.6 | 15970 | 8 | AQ839852 260L13-C5 |

| | | | | | |
|----|-------|------|------|---|--------------------|
| 25 | 211.4 | 10.6 | 537 | 8 | AQ390925 CITR1-E1- |
| 26 | 211.4 | 10.6 | 634 | 9 | AG019812 Homo sapi |
| 27 | 211.4 | 10.6 | 654 | 7 | CK003802 AGENCOURT |
| 28 | 211.4 | 10.6 | 687 | 6 | CD237038 ENPAPH09 |
| 29 | 211.4 | 10.6 | 692 | 5 | B0662309 C184C05.2 |
| 30 | 211.2 | 10.6 | 338 | 2 | AM502796 UI-HF-BR0 |
| 31 | 211.2 | 10.6 | 596 | 8 | AQ320557 RPTC11-96 |
| 32 | 211.2 | 10.6 | 801 | 8 | BZ600703 WHAC76TF |
| 33 | 211 | 10.5 | 552 | 8 | B68316 CIT978SK-A- |
| 34 | 210.8 | 10.5 | 657 | 7 | CF123536 UI-HF-CH0 |
| 35 | 210.6 | 10.5 | 1016 | 5 | BQ674146 AGENCOURT |
| 36 | 210.4 | 10.5 | 388 | 2 | AM069227 C141H09.x |
| 37 | 210.4 | 10.5 | 735 | 7 | CF146929 UI-HF-CB0 |
| 38 | 210.2 | 10.5 | 671 | 9 | AG079103 Pan trogl |
| 39 | 210.2 | 10.5 | 849 | 6 | CB991080 AGENCOURT |
| 40 | 209.8 | 10.5 | 397 | 1 | A1421950 C145D05.x |
| 41 | 209.8 | 10.5 | 463 | 2 | AM026305 wv11d10.x |
| 42 | 209.6 | 10.5 | 461 | 8 | B65075 CIT-HSP-201 |
| 43 | 209.6 | 10.5 | 501 | 5 | BX485916 DKFZP686E |
| 44 | 209.6 | 10.5 | 604 | 5 | BM990713 UI-H-D10- |
| 45 | 209.2 | 10.5 | 534 | 6 | CB050438 NISC_gj17 |

ALIGNMENTS

RESULT 1
CD245375/c
LOCUS
DEFINITION
AGENCOURT 14128276 NIH MGC 181 Homo sapiens cDNA clone
IMAGE:30374350 5', mRNA sequence.
CD245375
VERSION
CD245375.1 GI:31005839
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 908)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTE OF HEALTH
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
Plate: NDAM433 row: j column: 23
High quality sequence start: 10
High quality sequence stop: 557.
Location/Qualifiers
1. 908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30374350"
/feature_type="white matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistancees) "
/note="Vector: pCMV-Sport6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (Rcovy site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 11.3%; Score 227; DB 6; Length 908;
Best Local Similarity 69.2%; Pred. No. 9.9e-26;

Matches 402; Conservative 0; Mismatches 160; Indels 19; Gaps 6;

```
Qy 437 ACATTAAGAGGTAAGGCTGGGCGAGTGGCTCACACCTGTAATCCGACACTT 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 680 AAATGTAAGAAAAAGTGGCGGCGGCGGCGGCTGATCCGACACTT 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 497 GGGAGGCTGAGGCGAGGATCACTTTGGTCAAGAGTTTGAAGTACCTGGCCAAAT 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 GGGAGGACGAGGCGGCGGATCAC--GAGCTCGAAGATCGAAGCCATCTGTTAAAC 563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 557 GGGGAACTGCTGCTCAAAAAAATTAATAAAATTAAGCTGGATGTGGCGGCTT 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 GGGGAACTGCTGCTCAAAAAAATTAATAAAATTAAGCTGGATGTGGCGGCTT 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 617 GTATCTCAGCTGCTCAGAGCGGAGACAAGAAATCACTTAAACCCAGAGGTGAGG 676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 GTAGTCCAGCTACTTGGAGGCTGAGGCGAGAGAAATGGGTGAACCCGAGGACAGC 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 677 TTGCACTGAGCTGAGATGCGCACTGCACTCAACCTGGAGACAGAGTGAACCTTTGT 726
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 TTGCACTGAGCTGAGATGAGACCACTGCACTCCACGCTGGGCGAGAGACAG- ACT 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 737 CTCAAAAAAGAAAAAACAAGTAATAAAAGAAAGAGTGAAGTTAACTTAAATACC 796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 383 CTGCTCAGAAAAAAGAAAAAAGAAATAGTCAATTTTAAATTAATTTAGCTC 324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 797 AATGATCCCAATATACATTTCAAGTATTAATTAATAAACAATTATGAGTGA 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 323 AGATGCTCAAAATATATATATTTAA-----CATGATACCAAAAGTTATTAACAAA 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 857 TACTTACATTTCTTTCTTTTCAATTAATTAAGCTTTGAAAGTATATA- TGTAT 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 TATCTCAATTCCTGTTGG--GGGTATTAATCTTTGAAATTAATTAATTTTAT 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 916 GCTGACGACATCTCAATTTGAGCTAGCTACATTTAGCTGCTCACTACCAATGTG 975
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 ACTTACAGCTGATCTTATTCAGATGACCAATTTCCAGTGTCTCACTAGCCCATGTGG 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 976 CTAGCA-----GTTACTGATTGATGACGAGATCTAGA 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 CTAGCACTATGTTACTATCTGATAGCAGAGCTATTA 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 2
BZ603010 694 bp DNA linear GSS 08-JUN-2003
LOCUS BZ603010
DEFINITION WHAA35TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
sapiens genomic clone MCF7_1-6F22, genomic survey sequence.
ACCESSION BZ603010
VERSION BZ603010.1 GI:31511472
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 694)
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
JOURNAL MEDLINE
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
USCF Comprehensive Cancer Center
USCF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..694

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-6F22"
/sex="female"
/clone_1ib="Human MCF7 breast cancer cell line library
(MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."

ORIGIN

Query Match 11.1%; Score 223; DB 8; Length 694;
Best Local Similarity 79.1%; Pred. No. 4,5e-25;
Matches 265; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
Qy 439 ATTAAACAGTAAAGGCTGGGCGAGTGGCTCACACTGTAATCCGACACTTTGG 498
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28 ACTAAAGATGAAATGTTGCTGGGCGGCTGGCTCAGGCTGTATCCAGACACTTTGG 87
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 499 GAGGCTGAGGCGAGGATCACTTTGGTCAAGAGTTTGAACACTGCTGGCCAAATGG 558
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 GAGGCTGAGTGGGCGAGTATGCTGAGTCAAGAGTTCCAGACAGCCTGGGCAACATGG 147
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 559 CGAAACTGCTCTTAAAGAAAAATTAATAAAATTAAGCCGTGAGTGGCGGCGCTGT 618
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 148 CGAAACCTGTCTTAAAGAAAAATTAATAAAATTAAGCCGTGAGTGGCGGCGCTGT 207
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 619 AATTCAGCTGCTCAGAGGCGGAGACAGAAATCACTTAAACCCAGAGGTGAGATT 678
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 AATCAAGCTACTCGGAGGCTGAGGCAAGAAATGCTTTGAGCCAGGAGCAAGGTT 267
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 679 GCAGTGAAGTGAATGCTGCCACTCACTCAACCTGGAGACAGAGTGAACATTTGTCT 738
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 GCAGTGAAGTGAATGCTGCCACTCACTCAACCTGGAGACAGAGTGAACATTTGTCT 327
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 739 CAAAAAGAAAAAACAAGTAATAAAAGAAACA 773
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 362
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AF119908 1750 bp mRNA linear HTC 08-MAY-2001
LOCUS AF119908
DEFINITION Homo sapiens PRO2955 mRNA, complete cds.
ACCESSION AF119908
VERSION AF119908.1 GI:7770252
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1750)
Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
Functional prediction of the coding sequences of 79 new genes
deduced by analysis of cDNA clones from human fetal liver
Unpublished
2 (bases 1 to 1750)
Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M.
and He,F.
Direct Submission
Submitted (13-JAN-1999) Department of Experimental Hematology,
Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing,
Beijing 100850, P. R. China
location/Qualifiers
1..1750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```
/clone="FLC0974"
/tissue_type="liver"
/dev_stage="fetus"
379..677
/rpc_family="Alu"
/rpc_type=dispersed
/complement(1163..1396)
/notes="predicted protein of HQ2955"

CDS
repeats_region

ORIGIN
Query Match      11.1%; Score 221.6; DB 3; Length 1750;
Best Local Similarity 80.7%; Pred. No. 5.7e-25;
Matches 271; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 437 ACATTAAACAGTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 496
    |||||
DB 359 ACACGAAGGTGGGGAATAGCCGGCGCAGTGGCTCACGCCCTGTATCCAGCACTTT 418
    |||||
QY 497 GGGAGGCTGAGGCGAGATCACTTTGTCTAGAGATTGAGCTAGCCCTGGCCACAT 556
    |||||
DB 419 GGGAGGCGGAGGCGAGATCACTTGAAGATTGAGACCAAGCTGACCAACAT 478
    |||||
QY 557 GGGCAAACTGCTCTTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 616
    |||||
DB 479 GGTAAACCTGCTCTTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 538
    |||||
QY 617 GTAAATCTAGTCTTCAAGAGCGGAGACACAAATCACTTAAACCGAGAGGTGAGG 676
    |||||
DB 539 GTAAATCCAGTACTCAGAGGCTGAGGAGGAGAAATCACTTGAACCGGAGGAGGAGG 598
    |||||
QY 677 TTGAGAGAGCTGAGTCCGCGCACT-CAGTCCAACTGGGAGACAGATGACCTTTTG 735
    |||||
DB 599 TTGCAAGTGAAGTGAATCAATGCGCTGCACTTCAAGCTGGGGGACAGAGAGACTGTGT 658
    |||||
QY 736 TCTCAAAAAGAAAAAACAAGTAAAAAGAAA 771
    |||||
DB 659 CTCAAAAAGAAAAAAGATGAGAAATTAATCA 694
    |||||

RESULT 4
CD370173      769 bp      mRNA      linear      EST 05-AUG-2004
LOCUS         UT-H-FT1-bkb-e-05-0-UI.s1 NCI_CGAP_FTI Homo sapiens cDNA clone
DEFINITION    UT-H-FT1-bkb-e-05-0-UI 3', mRNA sequence.
ACCESSION     CD370173
VERSION       CD370173.1 GI:31154263
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 769)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgabs-rt@mail.nih.gov
              Tissue Procurement: Dr. Gary W. Hunninghake, U of I
              CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
              CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Distribution information can be found at
              http://genome.uiowa.edu/distribution/cgap.html
              The following repetitive elements were found in this cDNA
              sequence: 346-636, >Alu
              Seq primer: M13 FORWARD
```

```
FEATURES
POLYA=Yes.
Location/Qualifiers
1..769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT1-bkb-e-05-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Lung; Vector: pRT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_FTI is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The RNA samples
were a mixture of these conditions (lines refer to
incubations following isolation by bronchoalveolar lavage)
(some normal donor macrophages were cultured in some of
the conditions, other donor macrophages in different
conditions). The mRNA samples were pooled for library
construction. Control 0 hours; control 3 hours; control 24
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus
moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral
vector (Ad5 CMV egfp), moi 500, 3 hours; Adenoviral vector
(Ad5 CMV egfp), moi 500, 24 hours; wt adenovirus moi 500,
3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS
3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3
hours; wt adenovirus + LPS 24 hours. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pRT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GGCCATGCGG. The tissue was provided by Dr. Gary W.
Hunninghake of the University of Iowa.
TAG TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UT-H-FT1
TAG_SEQ=GGCCATGCGG"

ORIGIN
Query Match      11.0%; Score 220; DB 6; Length 769;
Best Local Similarity 80.4%; Pred. No. 1.3e-24;
Matches 270; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 437 ACATTAAACAGTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 496
    |||||
DB 326 ACACGAAGGTGGGGAATAGCCGGCGCAGTGGCTCACGCCCTGTATCCAGCACTTT 385
    |||||
QY 497 GGGAGGCTGAGGCGAGATCACTTTGTCTAGAGATTGAGCTAGCCCTGGCCACAT 556
    |||||
DB 386 GGGAGGCGGAGGCGAGTGAATCACTGAGGCCGAGATTGAACCAAGCTGACCAACAT 445
    |||||
QY 557 GGGCAAACTGCTCTTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 616
    |||||
DB 446 GGTAAACCTGCTCTTAAAGGCTGGGCGCAGTGGCTCAACCTGTATCCAGCACTTT 505
    |||||
QY 617 GTAAATCTAGTCTTCAAGAGCGGAGACACAAATCACTTAAACCGAGAGGTGAGG 676
    |||||
DB 506 GTAAATCCAGTACTCAGAGGCTGAGGAGGAGAAATCACTTAAACCGGAGGAGAGG 565
    |||||
QY 677 TTGCAAGTGAAGTGAATCAATGCGCTGCACTTCAAGCTGGGGGACAGAGACTTTTG 735
    |||||
DB 566 TTGCAAGTGAAGTGAATCAATGCGCTGCACTTCAAGCTGGGGGACAGAGAGACTTTGT 625
    |||||
QY 736 TCTCAAAAAGAAAAAACAAGTAAAAAGAAA 771
    |||||
```

Db 626 CTCAAAAAAGATGAGAAATATCA 661

RESULT 5
BX491983 500 bp mRNA linear EST 04-SEP-2003
LOCUS DKEZp78100513.1 781 (synonym: h1cc4) Homo sapiens cDNA clone
DEFINITION DKEZp78100513.5', mRNA sequence.
ACCESSION BX491983
VERSION BX491983.1 GI:32002977
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Oesanger,A.,
Fodor,G., Han,W. and Wiemann,S.
TITLE EST (Koehler,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Heinrich-
Heine-University, Dueseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No 3' sequence available.
This clone (DKEZp78100513) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: source@rzpd.de.

FEATURES
source
1..500
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKEZp78100513"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="781 (synonym: h1cc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI; cDNA-collection"

ORIGIN

Query Match 11.0%; Score 219.6; DB 5; Length 500;
Best Local Similarity 82.7%; Pred. No. 1.7e-24;
Matches 263; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

QY 443 AAACAGTAAAGAGCTGGCGGCGAGTGCACACCTGTATCCAGCACTTGGAGG 502
Db 13 ATAAAGAAATAGAGCTGGGCACTGAGGCTCAAGCTTATCCAGCACTTGGAGG 72
QY 503 CTGAGGCAAGCAATACCTTTGGTCTGAGAGTTGAGACTAGCCTGCCAATGGCGAA 562
Db 73 CCGAGGGGGGCGAGTCACTGAGGCTCAGGCGTTGAGACTAGCCTGCCAATGGCGAA 132
QY 563 ACTGTGCTCTAAATAAATAAATAAATTAGCTGTCATGTCGAGGGGCGCTGTATC 622
Db 133 ACCCATCTTTACTTAAATAAATGCAAAATTAAGCGGGGTGTGTGAGCGCTGTATC 192
QY 623 TCAAGTCTCAGAGGCGGAGACACAAGAAATCACTTAAACCCAGAGAGTGAAGTTGAG 682
Db 193 CAGCTACTTGGAGGCTGAAAGCAGAGAAATCGTTAAGTAGAGAGTGAAGTTGAG 252
QY 683 TGAAGTGAATCGTCCACT-CACTCAACCTGGAGACAGAGTGAACATTTTGTCTCA 741
Db 253 TGAGCTGAGATTGGCCACTGCACTTAAGCCTGGGTGACAGAGTGAACATTTTGTCTCA 312
QY 742 AAGAAAAAATCAAA 759
Db 313 AAAAAAATCAAA 330

RESULT 6
BE790769/c 837 bp mRNA linear EST 20-SEP-2000
LOCUS BE790769
DEFINITION BE790769.1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936325 5',
mRNA sequence.
ACCESSION BE790769
VERSION BE790769.1 GI:10211967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW780 row: c column: 14
High quality sequence stop: 723.

FEATURES

source

1..837
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936325"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC_7"
/note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 10.9%; Score 218.4; DB 2; Length 837;
Best Local Similarity 70.0%; Pred. No. 2.2e-24;
Matches 308; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 332 AAGACAGTGAAGACCCCTGTCTCAAGAGTCATCCCTAGAGTGTGTAGTGTAGT 391
Db 472 AAAAAAGTAAAGCCCTTTATGGGCTGAATCATGTCTGCTGAATTCCTATGTTGAAG 413
QY 392 AATAATGCTGAGCTGCTTATGATCATTTCCAGTTTATAGTCCACATTAAACAGTAA 451
Db 412 CCTTAACCCCGAGTACTGAGATGATGATTTGATATAGGAGCCCTTAAAGGAGTAA 353
QY 452 AAAAAAGCTGGGCGAGTGGCTCAACCTGTATCCAGCACTTTGGAGGCTGAGGCGAG 511
Db 352 ATTAAAGCTGGGTGAGTGGCTCACACCTATATCCATCACTTTGGAGGCGGAGGAG 293
QY 512 GCAGATCACTTTGTGTCAGAGTTTGAAGTGTAGCCCTGGGCAATGAGGAGAACTGTCT 571
Db 292 GGTATACCTTGAAGTCAAGAGTTTGAAGCAGCTGACCAATTAATAAACCCTGTCT 233
QY 572 CTAAAAAATAAATAAATAAATTAGCTGATGATGGGCGGCGCTGTATATCTCAAGTGTCT 631
Db 232 CTACTTAAAAAATCAAAATTTAGCTTGTGTGTGTATACCGCTGTATATCTCAAGCACT 173
QY 632 CAGAGGCGGAGACACAAAGATCACTTAACCCAGAGAGTGAAGTGTGAGTGAAGTGA 691

Db 172 TGGAGGCTAAGCTGTGAGAAATCACTGAAACCCGGAGACAGAGTTGTCAGTGAAGTGTG 113

Qy 692 ATCGGCGCACT-CACTGCAACCTGGGAGACAGAGTGAACATTTGTCTCAAAAAGAAA 750

Db 112 ATCGGCGCACTGCACTGAGCTGGGTGACAGAGGAGACTGTGTCTCAAAAAGAAA 53

Qy 751 AAAAAACAAGTAAAAAGAA 770

Db 52 AAAAAAAGAGTAAATTA 33

RESULT 7
B71494/c
LOCUS
DEFINITION B71494 649 bp DNA linear GSS 08-APR-1999
Rpci11-8N9 TP Rpci-11 Homo sapiens genomic clone Rpci-11-8N9,
genomic survey sequence.
ACCESSION B71494
VERSION B71494.1 GI:2710715
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Adams,M.D., Rounalev,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Sub,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: Rpci11-8N9_TV
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library Rpci-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.butlalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.butlalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends

FEATURES
Location/Qualifiers
1..649
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7503008"
/db_xref="taxon:9606"
/clone="Rpci-11-8N9"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="Rpci-11"
/note="Vector: pBACe3.6, Site 1: EcoRI, Site 2: EcoRI,
Rpci11 Human Male BAC library"

ORIGIN
Query Match 10.8%; Score 216; DB 8; Length 649;
Best Local Similarity 78.2%; Pred. No. 5.6e-24;
Matches 297; Conservative 0; Mismatches 80; Indels 3; Gaps 3;

Db 412 TGTGATTTCCAGTTTATTAAGCCACATTAACAGGTAAAGAGCTGGCGCAGTGG 471

Db 553 TGTACTATAGGAAATATATGAGAGCTTTAAAGACCTTTTACAGCCCGGTGTGTGG 494

Qy 472 CTCACACCTGTAATCCAGACCTTTGGAGGCTGAGGACAGATACCTTTGTGTAGG 531

Db 493 CTTACACCTGTAATCCAGACCTTTGGAGGCTGAGGACAGATACCTTTGTGTAGG 434

Qy 532 AGTTGAGACTAGCGCTGGCCAAACATGGCGAAACTCTGTCTTAAAAAAATACAAAT 591

Db 433 AGTTGAGACAGAGCTGGCCCAACATGSCAAATCTGTCTCT-ACATAAATAACAAAAT 375

Qy 592 TAGGCTGCAATGAGGCGGCGCCTGTAAATCTAGCTGTCTCAGAGGCCGACACAGAA 651

Db 374 TAGCAGAGAGTGTGGGCGCATGCTTAATCCAGCTACTTAGAGGCTGAGGACAGAA 315

Qy 652 ATCACTTAACCCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 710

Db 314 ATTACTTGAACCTGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 255

Qy 711 CTTGGAGACAGAGTGAACCTTTGTCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAA 770

Db 254 CTTGGAGACAGAGTGAAC-TCGCTCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 196

Qy 771 ACAGGTGAAGTAACTTAA 790

Db 195 AACCTTCAGTAACTTAA 176

RESULT 8
CD518470
LOCUS
DEFINITION CD518470 1019 bp mRNA linear EST 06-JUN-2003
AGENCOURT_14368329 NIH MGC 181 Homo sapiens CDNA clone
IMAGE:30396254 5', mRNA sequence.
ACCESSION CD518470
VERSION CD518470.1 GI:31450188
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9abds-f@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDA477 row: k column: 15
High quality sequence stp: 558.

FEATURES
Location/Qualifiers
1..1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30396254"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 181"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (copy site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH MGC Library."

ORIGIN
Query Match 10.8%; Score 216; DB 6; Length 1019;
Best Local Similarity 73.2%; Pred. No. 4.9e-24;
Matches 287; Conservative 0; Mismatches 104; Indels 1; Gaps 1;

Qy 433 AGCCACATTAACAGTAAAGAGCTGGCGCAGTGTCTCACCTGTAATCCAGCA 492

Db 221 ACCCAGAGTAAGATTAAGAGGCGCCAGGTGCACTGCTGCTGTAATTCAGCA 280

| | | | |
|----|-----|---|-----|
| OY | 493 | CTTTGGAGGCTGAGGACGACATCACTTTGGTCAGAGTTTAACTAGCTGGGCA | 552 |
| Db | 281 | CTTTGGAGGCTTAGGCAAGTGATCACTTAGGTCAGAGTTTGAACACAGCTGGGCA | 340 |
| OY | 553 | ACATGGCAAACTCTGTCTCTTAAAAAAAATTCAAAAATTAGCTGGCATGGTGGCGGC | 612 |
| Db | 341 | ACATGGTAAAAACCCGGTCTT-ACATAAAAAATTCAAAAATTAGCTGGCATGGTGGCGGAC | 399 |
| OY | 613 | GCTGTAAATCTCAGCTGCTCAGAGAGCCGAGACACAAGATCACTTAAACCGAGAGGTG | 672 |
| Db | 400 | ACCTGTATCTCAGTTATTTCAGAGAGCTGAGGACAGAGATTCGTTGAGCCAGGAGGTG | 459 |
| OY | 673 | GAGGTTGCAGTGAAGTGAATCGTCCACTACTCAACTCTGGGAGACAGATGACATT | 732 |
| Db | 460 | GAGGTTGCAGTGAAGTGAATCAACCAATTGCATCAGCTGGGGGACAGGAGCAAACT | 519 |
| OY | 733 | TTGTCTCAAAAAGAAAAAAAACCAAGTAAAAAGAAACAGGTGAAGTTAACTTTAATA | 792 |
| Db | 520 | CTGTCTCAAAAANNN | 579 |
| OY | 793 | ACCCAATGTATCCCAATTCATCTTTCAA | 824 |
| Db | 580 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 611 |

| | |
|------------|---|
| RESULT 9 | |
| CR749224 | |
| LOCUS | 2449 bp mRNA linear HTC 19-AUG-2004 |
| DEFINITION | Homo sapiens mRNA; CDNA DKFZp781C2298 (from clone DKFZp781C2298). |

| | | |
|----------|----------------------|-------------|
| VERSION | CR749224.1 | GI:51476179 |
| KEYWORDS | HTC. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Homo sapiens | |

REFERENCE
AUTHORS

Ostenwælder, B., Obermaier, B., Deutschenaur, S., Schapp, A.,
1 (bases 1 to 2449)

Euryarchaea; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

| | |
|------------------|---|
| CONSRMT TITLE | The German cDNA Consortium |
| JOURNAL | Direct Submission |
| COMMENT | Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuhofberg, GERMANY Clone from S. Wemann, Molecular Genome Analysis, German Cancer |

Reese/Center (June 27). Email: B.Wedemund@k2-Net.de. de/ Sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German genome project. This clone (DKFZp781C2298) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: <http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp781C2298> Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

```

source
1..2449
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781C2298"
/tissue_type="colon carcinoma"
/clone_lib="781 (synonym: hicc4). Vector pSport1_Sfi; host
DH10B, sites SfiI + SfiIb"
/dev_stage="adult"
/note="hypothetical protein"
1..2449
/gene="DKFZp781C2298"
161..862
/gene="DKFZp781C2298"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18080.1"

```

```

/dbs xref="GI:51476180"
/translation="MLREATYKKSKEKPEMALPOGRILFPRDVAIEFSLEBEMCLNPA
ORLYALVYMLENTRNLEFPVNDISSKMMKBESSSAOENTVEITHGTLORHSHSLIDPC
FOELBRDINHFEFQMOEBERNGEAPWTIKETLQTSRDHDDQDHANKPIKIDLOGSSP
HSHMLPELHIFQPEMKIGNOVEKSIINNLSLITLITSCIPSKTINNNYGNNSLHSLUP
IQKLGSTHERKILPMP"

```

| | | | | |
|-----------------------|--------------|-------------|---------------|-------------|
| Query Match | 10.8% | Score 215.6 | DB 3 | Length 2449 |
| Best Local Similarity | 79.7% | Pred. No. 4 | 4e-24 | |
| Matches 279 | Conservative | 0 | Mismatches 69 | Indels 2 |
| | | | | Gaps 2 |

| Accession | Sequence | Position |
|-----------|--|----------|
| Oy | TTACTACCCGACATTAAGAAACAGGTATTAAGAAAAAGAGCTGGGCGCAGTGGCTCACACCTGTATCC | 487 |
| Db | TTAGTGTCTTATCTTTGAAAACATGTACTTACGCTGGGGCGGTGGCTCACGCCCTGTATCC | 2020 |
| Oy | CAGCAGCTTTGGAGAGGCTGAGGACAGCAGATCACCTTTGGTACGAGTTTGAAGACTAGCCT | 547 |
| Db | CAGCAGCTTTGGAGAGGCTGAGGTAGTGGATCACCTGAGGTACAGAGATTGTAGACCAAGCTT | 2080 |
| Oy | GGCCCAATGAGCGGAACTCTGTCTCTCTTAAAAAAATACAAAAATTAGCTCGGCAATGGTGG | 607 |
| Db | GACCAAAAAGGTCAAACTCCGTCTCT-ACCTAAAAATACAAAAATTAGCCCAAGATGGTAG | 2139 |
| Oy | CGGCGCCTGTATCTCAGCTGCTCAGAGAGCCGAGACACAGAAATCATTTAAACCCAGG | 667 |
| Db | CAGACGCGCTGTAGTTCCAGCTAATTCAGAGAGATGAGACAGAGAAATTGCTTGAACCCAGG | 2140 |
| Oy | AGGAGAGGTTGACGTAGCTGAGATGATGATCCACT-CATCCAAACCTGGGAGACAGAGT | 726 |
| Db | AGGTGAGGCTGACGTAAAGCCGAGATTTTGCACCTGCACCTTACCTGGGCAACAGAGT | 2259 |
| Oy | ACACTTTTGTCTCAAAAAGAAAAAAACAAAGTAAAAAGAAACAGGT | 776 |
| Db | AGACTCTTGTCTCCAAAAAGAAAAAGAAAGACAAAAAGAAACATGT | 2309 |

| | |
|-----------------|--|
| RESULT | 10 |
| LOCUS | BX477553 |
| DEFINITION | BX477553 573 bp mRNA linear EST 04-SEP-2003 |
| ACCESSION | DKRZPE86H1196.F1.686 (synonym: hlcC3) Homo sapiens cDNA clone |
| VERSION | DKRZPE86H1196_5', mRNA sequence. |
| KEYWORDS | BX477553 BX477553.1 GI:31911608 |
| SOURCE | EST. |
| ORGANISM | Homo sapiens (human) |
| AUTHORS | Homo sapiens Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euarchontomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 573) |
| TITLE | Mammut.R., Heubner.D., Mwes.H.W., Well.B., Amlid.C., Osanger.A., Robo.G., Han.M. and Wiemann.S. |
| JOURNAL COMMENT | EST (Mammut.R., Heubner.D., Mwes.H.W., Well.B., Amlid.C., et al.) Unpublished (2003) Contact : MIPS |

```

FEATURES
source
    1..573
        /organism="Homo sapiens"
        /mol_type="RNA"
        /db_xref="taxon:9606"
        /clone="DKFZp686H1196"
        /dev_stage="adult"
        location/Qualifiers
            No sl sequence available.
            This clone (DKFZp686H1196) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heuberger 6, 14059
            Berlin-Charlottenburg, Germany; Email: clone@rzpd.de.

```


ORIGIN

/lab host="DH10B"
/clone_lib="686 (synonym: h1cc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

Query Match 10.8%; Score 215.4; DB 5; Length 573;
Best Local Similarity 82.0%; Pred. No. 7.3e-24;
Matches 260; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 464 CGCAGTGGCTCACCTGTAAATCCAGCACTTTGGAGGCTGAGCGACGAGATCACCTT 523
DB 1 CGCAGTGGCTCACCTGTAAATCCAGCACTTTGGAGGCGAGGAGGAGTGAATCTTG 60
QY 524 TGGTCAGAGCTTAAAGCTAGCTGGCCCAATGGCCAAATCTGTCTCTTAAATAAAT 583
DB 61 AGGTGAGAGTTCAAGACTAGCTGGCCCAATGGCCAAATCTGTCTCTTAAATAA 120
QY 584 ACAAAATTAAGCTGAGTGGTGGCGGCGCTGTAAATCTGAGTGTGCTGAGAGCGGAG 643
DB 121 AAAAAATTAAGCGCATGTGGCGGCACTGTAAATCCAGCGCTCGGAGGCTGAA 180
QY 644 ACACAGAAATCACTTAAACCCAGAGGTGAGGTTGAGTGAATGATGTCCTACT- 702
DB 181 GCAGAGAAATCGCTTGAACCCAGAGGAGGTTGAGTGAATGATGTCCTGACTG 240
QY 703 CACTTCAACTGGAGACAGAGTACCTTTGTCTCAAAAAAGAAAAAACAAGTA 762
DB 241 CACTTCAAGCTGGGCGACAGATCGACTGTCTCAAAAAAGAAAAAAGAAAAAAGA 300
QY 763 AAAAAAGAACAGGTGAA 779
DB 301 GAAAAACGAGGAGAAAGA 317

RESULT 11 823 bp mRNA linear EST 15-MAY-2003
CD101954/c
LOCUS
DEFINITION
AGENCOURT 13980468 NIH_MGC_187 Homo sapiens cDNA clone
IMAGE:30373404 5', mRNA sequence.

ACCESSION
CD101954
VERSION
CD101954.1 GI:30755128
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 823)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: gsapbs-j@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDCM167 row: c column: 13
High quality sequence stop: 327.
Location/Qualifiers
1..823

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30373404"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_187"
/note="Organ: Blood vessels - aorta, basilar and artery;
Vector: pDNR-LIB; Site 1: SfiI (ggccataggc); Site 2:
SfiI (ggcgctcgcc); 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:
5'-CAGGCGCATTAAGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGGCCAGAGCGGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.5-4.0 kb). 14/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 10.8%; Score 215.4; DB 6; Length 823;
Best Local Similarity 80.9%; Pred. No. 6.5e-24;
Matches 263; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

QY 444 AACAGTAAAAAAGGCTGGCGGCGAGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGC 503
DB 326 AAGAAATTAATTAAGGCGGCGAGTGAAGTGGCTCAACCTGTAAATCCAGCACTTTGGAGGC 267
QY 504 TGAAGCAGAGATCACTTGTGTCAGAGATTGAGACTAGCTGGCCCAATGCGCAATGCGGAA 563
DB 266 TGAAGCAGAGATCACTTGTGTCAGAGATTGAGACTAGCTGGCCCAATGCGCAATGCGGAA 207
QY 564 CTGTGTCTTAAAAAATAAATAAATTAAGCTGGCATGTGCGGCGCTGTAAATCT 623
DB 206 CCTGTCTCT-ACATAAATAAATAAATAAATTAAGCTGTGTCATGTGCGGCGCTGTAAATCT 148
QY 624 CAGCTGCTCAGAGGCGGAGACCAAGATCACTTAAACCCAGAGGTGAGGTTGCACT 683
DB 147 CTGTCTTGTGGAGGCTGAGACCAAGATCACTTGAACCCAGAGGTGAGGTTGCACT 88
QY 684 GAGCTGAGATCGGCGCACTCACTCCAACTGGGAGCAGAGTGACCTTTGTCTCAAA 743
DB 87 GAGCTGAGATCAATCCACTGCACTAGCTGGGCGACAGCAAAACTGTCTCAAAA 28
QY 744 AGAAAAAATAAATAAATAAATAAAG 768
DB 27 AAAAAAAAAAAAAAAAAAATG 3

RESULT 12 397 bp mRNA linear EST 19-OCT-2000
AV759632
LOCUS
DEFINITION
AV759632 MDS Homo sapiens cDNA clone MDSDBF02 5', mRNA sequence.
ACCESSION
AV759632.1 GI:10917480
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 397)
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Zeng, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA MDS clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..397

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSDBF02"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"

```

/1ab host="BM25.8"
/clone lib="MDS"
/note="Vector: pTriplEx2, Site_1: sflra, Site_2: sflrb

```

| | | | | |
|-----------------------|--------|---|-------|---------------------------------|
| Query Match | 10.7% | Score 215; | DB 1; | Length 397; |
| Best Local Similarity | 80.1%; | Pred. No. 9.4e-24; | | |
| Matches | 265; | Conservative | 0; | Mismatches 65; Indels 1; Gaps 1 |
| Qy | 458 | GCTGGCGCAGTGGCTCAACCTGTATATCCAGCATTTTGGAGGCTGAGCAGCAGAT | 517 | |
| Db | 1 | GCCGGTCCAGTGGCTCAACGCTGTATATCCAGCATTTTGGAGGCCAGGCGAGTGGAT | 60 | |
| Qy | 518 | CACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCAAACATGCGGAAACTGTGCTCTAAAA | 577 | |
| Db | 61 | CACCTGAGGTGAGGAGTTTGAGACCAAGCCTGGCCAAACATGTGGAAACCCGCTCTACTA | 120 | |
| Qy | 578 | AAAAATACAAAATTTAGCTGGCATGTGGGCGGGGCGCTGTATATCTCAGCTCTAAGAG | 637 | |
| Db | 121 | AAAAATACAAAATTTAGCTGGCCATGTGGGCGCTGTACCTGTATATCTTAGCTACAGAG | 180 | |
| Qy | 638 | GCCGAGACACAGAATCACTTTAAACCCAGAGGTGAGAGTTGCAGTGAAGCTAGATCGTG | 697 | |
| Db | 181 | GCTGAGGAGAGGAGATCCGCTTGAACTCTGGAGGTGTGTGTTGCAATGAGCCAAAGTACAA | 240 | |
| Qy | 698 | CCACT - CACTCCAACTTGGAGACAGAGTGACATTTTGTCTCAAAAAAGAAAAA | 756 | |
| Db | 241 | CCATTGCACTCCAGCTGGGTGTGACAGATGAGACTGTCTTCAMAGAGAAAAA | 300 | |
| Qy | 757 | CAGTAAAAAAGAAACAGGTGAAGTTAACTT | 787 | |
| Db | 301 | AAAAAAATGCGCGCCGCAATTTGAGCTCACTT | 331 | |

| | | | | |
|------------|---|--------|------|------------------------|
| RESULT 13 | BX644875/c | | | |
| LOCUS | BX644875 | 587 bp | mRNA | linear EST 04-SEP-2001 |
| DEFINITION | DKFZp781D1236_r1 781 (synonym: hicc4) Homo sapiens cDNA clone | | | |
| ACCESSION | DKFZp781D1236 5', mRNA sequence. | | | |
| VERSION | BX644875 | | | |
| KEYWORDS | BX644875.1 GI:34479208 | | | |
| SOURCE | EST. | | | |
| ORGANISM | Homo sapiens (human) | | | |
| REFERENCE | Homo sapiens | | | |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | |
| TITLE | 1 (bases 1 to 587) | | | |
| JOURNAL | Wambutt, R., Heubner, D., Mewes, H.W., Weill, B., Amlid, C., Osanger, A., | | | |
| COMMENT | Poco, G., Han, M. and Wiemann, S. | | | |
| | EST (Wambutt, R., Heubner, D., Mewes, H.W., Weill, B., Amlid, C., et al.) | | | |
| | Unpublished (2003) | | | |
| | Contact: MIPS | | | |
| | MIPS | | | |
| | IngoIsaeder Landstr.1, D-85764 Neuherberg, Germany | | | |
| | This is the 5' sequence of the clone insert | | | |
| | Clone from S. Wiemann, Molecular Genome Analysis, German Cancer | | | |
| | Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de; | | | |
| | sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing | | | |
| | consortium of the German Genome Project. | | | |
| | No s1 sequence available. | | | |
| | This clone (DKFZp781D1236) is available at the RZPD in Berlin. | | | |
| | Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 | | | |
| | Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. | | | |
| | Location/Qualifiers | | | |
| | 1..587 | | | |

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP781D1236"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="781 (synonym: hicc4)"
/notes=vector: psport1 st1, site_1: st11, site_2: st11b

```

cdna-collection™

| Query Match | 10.7%; | Score 215; | DB 5; | Length 587; |
|-----------------------|--------|---|-------|----------------------------------|
| Best Local Similarity | 80.1%; | Pred. No. 8,3e-24; | | |
| Matches | 265; | Conservative | 0; | Mismatches 65; Indels 1; Gaps 1; |
| QY | 442 | AAAACAGTMAAAAAAGCTGGGCGCAGTGGCTCAACCTGTAAATCCAGACATTTGGGAG | 501 | |
| Db | 463 | AGAAAAGMAAGAGAGCTGGGTGTGGTGCTCAGCCTGTAAATCCAGATCTTTGAGAG | 404 | |
| QY | 502 | GCTGAGCGAGGCAATCACTTTTGTGTCAGAGTTTGACATAGCTCTGGCCAAATGCGGA | 561 | |
| Db | 403 | GCCAGGGGGGTGATCACTTGAAGTCAGGAGTTTGAGACCAAGCTGGCCAAATGCGGA | 344 | |
| QY | 562 | AAACTGTGCTTMAAAAAAATACAAAATTAAGCGCTGGCATGTGGCGGCGCCTGTAT | 621 | |
| Db | 343 | ATTCTCTCTTATCAAAAAAATACAAAATTAAGCAGGCAATGTGTGCGCATGCTGTAT | 284 | |
| QY | 622 | CTCAGCTGCTCAGAGGCGGACACAGAATCACTTAAACCCAGAGGTGAGATTGCA | 681 | |
| Db | 283 | CCGAGCTAATCTGGGAGGCTGAGCGCAGAGAAATCTTGAAACCGGAGAGCGGAGTTGCA | 224 | |
| QY | 682 | GTGAGCTGAGATCGTGCCACT-CACCTCAACCTGGGAGACGAGTGAACCTTTGTCTCA | 740 | |
| Db | 223 | GTGAGTTGAGATTGTGCTACTGTACTCCAGCGCTGGGACACGAGAGGACTTATCCCA | 164 | |
| QY | 741 | AAAAGMAAAAAAACAAGTMAAAAAA | 771 | |
| Db | 163 | AAAAAAGMAAAGMAAAGMAAAGMAA | 133 | |

```

RESULT 14
LOCUS      CK905714/c
DEFINITION CK905714 633 bp mRNA linear EST 11-MAR-2004
            i176bh09.x6 Human insulinoma Homo sapiens cDNA clone IMAGE:563692
            3', mRNA sequence.
ACCESSION  CK905714
VERSION     CK905714.1
KEYWORDS    GI:45367245
SOURCE      EST.
ORGANISM    Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 633)
            Melton,D., Meadows,A., Clifton,S., Hiller,L., Marra,M., Pape,D.,
            Wylie,T., Martin,J., Bistrain,A., Schmitt,M., Theising,B.,
            Riter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
            McCann,R., Cole,R., Tsagarisbivili,R., Williams,T., Jackson,Y. and
            Bowers,Y.
            WashU-Harvard Pancreas EST Project
            Unpublished (2000)
            Other ESTs: i176bh09.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            This read is a 3' RESEQUENCE of a previously sequenced pancreas
            clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Seq primer: -40BP from Gibco.
            Location/Qualifiers
                1..633

```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="IMAGE:5636992"
/tissue_type="insulinoma"
```

/lab/hosts/UDH10B (phage-resistant)"
 /clone.lib-"Human insulinoma"
 /note="Organ: pancreas; Vector: pbluescript SK-; Site_1:
 XhoI; Site_2: EcoRI; Constructed with lambda ZapII system
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
 pbluescript SK- by Dr. H. Inoue following the Washington
 University protocol
 (http://genomew.wustl.edu/est/lambda_protocol.shtml).
 Please contact Hiroshi Inoue, MD/PhD for further
 information on this library (Metabolism Division, Permutt
 Laboratory, Washington University School of Medicine, Box
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
 is a Washington University Pancreas EST project library."

ORIGIN

| Query Match | Similarity | 10.7% | Score 215 | DB 7 | Length 633 |
|-------------|------------|-------------------------|--|------------|------------|
| Best Local | Similarity | 83.4% | Pred. No. 8.1e-24 | | |
| Matches | 256 | Conservative | 0 | Mismatches | 50 |
| | | | | Indels | 1 |
| | | | | Gaps | 1 |
| Qy | 454 | AAAGGCTGGGCGCAGTGGCTCA | CACCTGTAAATCCAGACATTTGGAGAGCTAGGCAAGC | 513 | |
| Db | 308 | ACAGGCCGGGCGCACTGGCTCA | CGGCTGTAAATCCAGACTTTTGGGAGGCGGAGGCAAGT | 249 | |
| Qy | 514 | AGATCACTTTGGTCAAGAGTTTG | AAGCTAGCTAGCCCAACATGGCGAAACTCTGTCTCT | 573 | |
| Db | 248 | GGATCACTCTGGGTGAGAGTTTG | AAGCAGCTTGGCCCAACATGGTGTAAACCCCGTCTCT | 189 | |
| Qy | 574 | AAAAAAAAAATCAAAAATTAG | CCGCGGCAATGGTGGGGGGCGCGCTGTAAATCTCAGCTGCATCA | 633 | |
| Db | 188 | ACTAAAAAAAAATCAAAAATTAG | CTGGGGCAATGGTGGGGCTCACTGTAAATCTTAGCTACTCA | 129 | |
| Qy | 634 | GGAGGCCGAGACACAAGAATCA | CTTAAACCGACGAGGTGAGAGTTGCAGTGAAGTGAAT | 693 | |
| Db | 128 | GGAAGCTTAGGCGAAGAGAAAT | CGCTTGAATCTGGGAGGTGTGTGTGTTGCAGTGAAGCCAAAT | 69 | |
| Qy | 694 | CGTGGCACT-CACTTCAA | CTCTGGGAGACAGAGTGAACACTTTTGTCTCAAAAAAGAAAAAA | 752 | |
| Db | 68 | CACACCAATTTGACATTCAC | CTGAGGTGAGACAGAGTGAAGTCTGTCTCAAGAAAGAAAAAA | 9 | |
| Qy | 753 | AAAAAA 759 | | | |
| Db | 8 | AAAAAA 2 | | | |

| | | | |
|------------|--|--------|------------------------|
| | RESULT | 15 | |
| | AM516097/c | | |
| LOCUS | AM516097 | 491 bp | mRNA |
| DEFINITION | xte0f05.x1 NCI CGAP Utc2 Homo sapiens cDNA clone IMAGE:2790849 3' | | linear EST 03-MAR-2000 |
| ACCESSION | similar to containe Alu repetitive element;; mRNA sequence. | | |
| VERSION | AM516097 | | |
| KEYWORDS | AM516097.1 GI:7154179 | | |
| SOURCE | EST. | | |
| ORGANISM | Homo sapiens (human) | | |
| | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 491) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: gcgaps-remai.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov/image/html/freesources.shtml Seq primer:-40UP from Gibco High quality sequence stop:424. | | |

FEATURES

Source

| FEATURES | SOURCE | Location/Qualifiers |
|----------|--------|--|
| 1. | .491 | |
| | | /organism="Homo sapiens" |
| | | /mol_type="mRNA" |
| | | /db_xref="taxon:9606" |
| | | /clone="IMAGE:2790849" |
| | | /tissue_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors" |
| | | /lab_host="DH10B" |
| | | /clone_11b="NCI_CGAP_Ut2" |
| | | /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr |
| | | Average insert 1.85 kb. Life Technologies catalog # 11539-012" |

ORIGIN

| Query Match | 10.7% | Score 214 | DB 2 | Length 491 |
|-----------------------|----------------|--|----------|------------|
| Best Local Similarity | 79.4% | Pred. No. 1,38-23 | | |
| Matches 278 | Conservative 0 | Mismatches 70 | Indels 2 | Gaps 2 |
| QY | 428 | TTAGTAGCCACATTAAAAACAGTAAAAAAGGCTGGGCGCAGTGGCTCAACACTGTATCC | 487 | |
| Db | 481 | TTAGTGTATTACTTGAAAAACATGATCTTAGGCTGGGCGGCGGTGCTCAAGCTGTATCC | 422 | |
| QY | 488 | CAGACATTTTGGGAGGCTGAGGCAAGCATCACTTTTGTACAGAGTTTGACTAGCTCT | 547 | |
| Db | 421 | CAGACATTTTGGGAGGCTGAGGTAGTGGATCACTGAGGTCAAGAGTTTGGACACAGCTT | 362 | |
| QY | 548 | GGCAACATGAGCGGAACTCTGTCTCTAAAAAAAATACAAAAATTAGCGTGGCATGGTG | 607 | |
| Db | 361 | GACCAAAAAGGTCAAACTTCGCTCTCT-ACAAAAATACAAAAATTGGCCACAGCATGGTG | 303 | |
| QY | 608 | CGGCGCGCTGTAACTTCAGCTGCTCAGAGGCGCAGACACAAAGATCACTTAAACCCAGG | 667 | |
| Db | 302 | CAGACGCTGTAGTTCCAGCTATTCAAGAGGATGAGACAGGAAATTGCTTGAACCCAGG | 243 | |
| QY | 668 | AGGTGGAGTTTGCAGTGAAGCTGAATGTGCACT-CACTTCAACTTGGGAGACAGAGTG | 726 | |
| Db | 242 | AGGTGGAGGCTGCAGTAAAGCGAGATCTTTCACATGCACTCTACCTGGGCAACAGAGTG | 183 | |
| QY | 727 | ACACTTTTGTCTCAAAAAGAAAAAAAACAGTAAAAAAGAAACAGGT | 776 | |
| Db | 182 | AGACTCTTGTCTCAAAAAGAAAAAGAAAAAAGACAAAGAAACATGT | 133 | |

```
Search completed: February 27, 2005, 12:42:21
Job time : 6279.68 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 ; Search time 795.397 Seconds

(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271b-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 agggagagctcggggaatga.....ccagaataactggaag 1571

Scoring table: IDENTITY_NUC
gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: N_Geneseq_16Dec04:*
1: geneseqn1980s:*
1: geneseqn1990s:*
1: geneseqn2000s:*
4: geneseqn2001as:*
4: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1571 | 100.0 | 11570 | 3 | AAA5905 Human KLK |
| 2 | 1571 | 100.0 | 11570 | 12 | ADK52482 Human kal |
| 3 | 1571 | 100.0 | 11570 | 13 | ADR72623 Human ren |
| 4 | 1571 | 100.0 | 11570 | 13 | ADR72875 Human ova |
| 5 | 541 | 34.4 | 567 | 12 | ACH74409 Human gen |
| 6 | 482 | 30.7 | 586 | 6 | ABL63579 Breast ca |
| 7 | 482 | 30.7 | 586 | 6 | ABL63596 Breast ca |
| 8 | 482 | 30.7 | 735 | 3 | AAC79469 CDNA sequ |
| 9 | 482 | 30.7 | 735 | 6 | ABK29013 Human bre |
| 10 | 482 | 30.7 | 738 | 2 | AAK84240 Human bre |
| 11 | 482 | 30.7 | 1260 | 10 | ADK80489 DNA encod |
| 12 | 482 | 30.7 | 1260 | 11 | ADN39197 Ovarian c |
| 13 | 482 | 30.7 | 1260 | 13 | ADR25550 Cancer/an |
| 14 | 482 | 30.7 | 1302 | 2 | AAK57990 Human BS2 |
| 15 | 482 | 30.7 | 1302 | 2 | AAK57989 Human BS2 |
| 16 | 482 | 30.7 | 1381 | 2 | AAZ06259 Human sec |
| 17 | 482 | 30.7 | 1381 | 8 | ADA40370 Human sec |
| 18 | 482 | 30.7 | 1387 | 10 | ADA56535 Gene enco |
| 19 | 482 | 30.7 | 1387 | 13 | ADR72622 Human ren |
| 20 | 482 | 30.7 | 1387 | 13 | ADR72874 Human ova |

| | | | | | |
|----|-----|------|------|----|---------------------|
| 21 | 482 | 30.7 | 1439 | 2 | AAZ06260 Human sec |
| 22 | 482 | 30.7 | 1439 | 8 | ADA40371 Human sec |
| 23 | 482 | 30.7 | 1439 | 10 | ADA56536 Gene enco |
| 24 | 482 | 30.7 | 1441 | 8 | ADA40368 Human sec |
| 25 | 482 | 30.7 | 1441 | 10 | ADA56533 Gene enco |
| 26 | 482 | 30.7 | 1481 | 12 | ADQ22212 Human 80f |
| 27 | 482 | 30.7 | 1490 | 8 | ADA40367 Human sec |
| 28 | 482 | 30.7 | 1490 | 10 | ADA56532 Gene enco |
| 29 | 482 | 30.7 | 1499 | 2 | AAK60578 Human ker |
| 30 | 482 | 30.7 | 1516 | 2 | AAZ06245 Human sec |
| 31 | 482 | 30.7 | 1516 | 2 | ADA40369 Human sec |
| 32 | 482 | 30.7 | 1516 | 10 | ADA56534 Gene enco |
| 33 | 482 | 30.7 | 1570 | 3 | AAZ65070 Membrane- |
| 34 | 482 | 30.7 | 1570 | 4 | AAK521471 Human CDN |
| 35 | 482 | 30.7 | 1570 | 5 | AAK44216 Human PRO |
| 36 | 482 | 30.7 | 1570 | 6 | ABK28605 Human DNA |
| 37 | 482 | 30.7 | 1570 | 8 | ACA64363 Novel hum |
| 38 | 482 | 30.7 | 1570 | 8 | ACA03830 CDNA enco |
| 39 | 482 | 30.7 | 1570 | 8 | ABK89368 DNA enco |
| 40 | 482 | 30.7 | 1570 | 8 | ABX80822 Human sec |
| 41 | 482 | 30.7 | 1570 | 8 | ACD44331 CDNA enco |
| 42 | 482 | 30.7 | 1570 | 8 | ACD42022 Human sec |
| 43 | 482 | 30.7 | 1570 | 8 | ABK79502 Human sec |
| 44 | 482 | 30.7 | 1570 | 8 | ACA93523 Novel hum |
| 45 | 482 | 30.7 | 1570 | 8 | ABK81205 Human sec |

ALIGNMENTS

RESULT 1
AAA5905
ID AAA5905 standard; DNA; 11570 BP.
AC AAA5905;
XX
DT 02-FEB-2001 (first entry)
XX
DE Human KLK-L2 gene.
XX
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KM kallikrein-like protein; serine protease; cytosolic; cancer;
XX
XX prostrate cancer; ds.
XX
OS Homo sapiens.
XX
XX
XX MO200053776-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000MO-CA000258.
XX
PR 11-MAR-1999; 99US-0124260P.
PR 01-APR-1999; 99US-0127386P.
PR 21-JUL-1999; 99US-0144919P.
XX
XX (MOUN) MOUNT SINAI HOSPITAL.
XX
XX Yousef GM, Diamandis EP;
XX WPI; 2000-587440/55.
XX P-PSDB; AAB21296.
XX
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
XX protein mediated disorders, especially cancer.
XX
PS Claim 1; Page 143-149; 184pp; English.
XX
CC The present sequence is the coding sequence of the human KLK-L2 gene,
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like
CC proteins are a subgroup of the serine protease enzyme family. They
CC catalyse the selective cleavage of specific polypeptide precursors to
CC release peptides with potent biological activity. Nucleic acids encoding

CC kallikrein-like proteins KLK-11, KLK-12, KLK-13, KLK-14, KLK-15 and KLK-16 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1571; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAGGAGGTATGGGGAATTGGAAGACAGAAAACAAATTAAGTCCAGCGAATGGATTCTTA 60
DB 10000 AAGGAGGTATGGGGAATTGGAAGACAGAAAACAAATTAAGTCCAGCGAATGGATTCTTA 10059
QY 61 TTGGAGGTATTTCTGCCCTAGAAACACTGGCAATACAGAGACATTTTGGTTGTCA 120
DB 10060 TTGGAGGTATTTCTGCCCTAGAAACACTGGCAATACAGAGACATTTTGGTTGTCA 10119
QY 121 CAACATATGAGGGGCAATTAAGTGAATAGATAGATGCAAGTGTCTGTCAAC 180
DB 10120 CAACATATGAGGGGCAATTAAGTGAATAGATAGATGCAAGTGTCTGTCAAC 10179
QY 181 ATGCTATGATGCAACAGGAGGCTCCACAAACCAATTATCAGATGCGCCAC 240
DB 10180 ATGCTATGATGCAACAGGAGGCTCCACAAACCAATTATCAGATGCGCCAC 10239
QY 241 AGTGCCCAATCGAGGAACCTCATCCAGGGCTGAAACCGTATTTTGGCAGAGGGAG 300
DB 10240 AGTGCCCAATCGAGGAACCTCATCCAGGGCTGAAACCGTATTTTGGCAGAGGGAG 10299
QY 301 GATTAAGATGGGTGTGGAGAAATGGGGAAGGAGTGTGTCCAGTAAGAAATAA 360
DB 10300 GATTAAGATGGGTGTGGAGAAATGGGGAAGGAGTGTGTCCAGTAAGAAATAA 10359
QY 361 GGGCTTGACAGAGCTGGAGGGGAGATGAGAGAGAAAGGAGGCGAGAGATACAGATGA 420
DB 10360 GGGCTTGACAGAGCTGGAGGGGAGATGAGAGAGAAAGGAGGCGAGAGATACAGATGA 10419
QY 421 GGGAGACAGGCTGGAGACAGAAAGTAGAGACGAAATTCAGATGTGAGAGAGAGGCTCA 480
DB 10420 GGGAGACAGGCTGGAGACAGAAAGTAGAGACGAAATTCAGATGTGAGAGAGAGGCTCA 10479
QY 481 CAGACCCGCCCAATGATGTGTGACCAACAGGAATCTGGAAGAGGAAGATGGAAGTGGAG 540
DB 10480 CAGACCCGCCCAATGATGTGTGACCAACAGGAATCTGGAAGAGGAAGATGGAAGTGGAG 10539
QY 541 AAGTGAAGAAATGGGCTTAAAGGTTGAATCTTGAGGCGCAGGCAATGTGCTCAAGCTCTTA 600
DB 10540 AAGTGAAGAAATGGGCTTAAAGGTTGAATCTTGAGGCGCAGGCAATGTGCTCAAGCTCTTA 10599
QY 601 ATCCCAACACTTTGAGGCTGAGAGTGGCGAATCACTTGAAGCCAGAGATTGAGACAG 660
DB 10600 ATCCCAACACTTTGAGGCTGAGAGTGGCGAATCACTTGAAGCCAGAGATTGAGACAG 10659
QY 661 CCGTGGCAACATGAGTGAACCCCGTCTTACAAAATAATPACAAAATAATTAAGCCGGGTGT 720
DB 10660 CCGTGGCAACATGAGTGAACCCCGTCTTACAAAATAATPACAAAATAATTAAGCCGGGTGT 10719
QY 721 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGCGAGAGAAATGCTTGAAC 780
DB 10720 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGCGAGAGAAATGCTTGAAC 10779
QY 781 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCACTGCGCTCAACTGGGCAACA 840
DB 10780 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCACTGCGCTCAACTGGGCAACA 10839
QY 841 GAGTAAAGCTCACTCAAAAAAATAAGCTGGAATTTGAGAGTAATTAATTAACATT 900
DB 10840 GAGTAAAGCTCACTCAAAAAAATAAGCTGGAATTTGAGAGTAATTAATTAACATT 10899
```

```
QY 901 CTCCTCTCTCTCTCTTTGGCTGTGTCTCATCTGTCTCTTTTGTGCAATTTCTTCATCT 960
DB 10900 CTCCTCTCTCTCTCTTTGGCTGTGTCTCATCTGTCTCTTTTGTGCAATTTCTTCATCT 10959
QY 961 CTGTACTTTTCATCTCTGTGTGTCTGTTCCTCATCTGTCTTTCTCATCTATGAGGATCTGTG 1020
DB 10960 CTGTACTTTTCATCTCTGTGTGTCTGTTCCTCATCTGTCTTTCTCATCTATGAGGATCTGTG 11019
QY 1021 GGTCTCTCATGTCTCTCTCTGCGCACTTGTGCAACATCTGTGCTCTCATGATCCGCCCTT 1080
DB 11020 GGTCTCTCATGTCTCTCTCTGCGCACTTGTGCAACATCTGTGCTCTCATGATCCGCCCTT 11079
QY 1081 TCTCTCTGACAGGATGATTTCTGGGGGCGCTGTGTCTGCAATGAGTCTCCCTGACGAGACTC 1140
DB 11080 TCTCTCTGACAGGATGATTTCTGGGGGCGCTGTGTCTGCAATGAGTCTCCCTGACGAGACTC 11139
QY 1141 GTGTCTGGGAGATTTACCTTTGTGCTCCGCGCCCAACAGACCGGGTGTCTACACGAACCTTC 1200
DB 11140 GTGTCTGGGAGATTTACCTTTGTGCTCCGCGCCCAACAGACCGGGTGTCTACACGAACCTTC 11199
QY 1201 TGCAGTTTCAACCAATGATCCAGGAAACATCCAGGCAACTCCTGAGTCAATCCAGGA 1260
DB 11200 TGCAGTTTCAACCAATGATCCAGGAAACATCCAGGCAACTCCTGAGTCAATCCAGGA 11259
QY 1261 CTCAGCAACCGGCATCCCACTTGTCTGACGAGGACAGGCTTGACACTCTCTTTCAGACCT 1320
DB 11260 CTCAGCAACCGGCATCCCACTTGTCTGACGAGGACAGGCTTGACACTCTCTTTCAGACCT 11319
QY 1321 CATTTCTTCCAGAGATGTTGAATGTTTCACTCTTCAGAGCCCTGACCCCATGTCTCT 1380
DB 11320 CATTTCTTCCAGAGATGTTGAATGTTTCACTCTTCAGAGCCCTGACCCCATGTCTCTCT 11379
QY 1381 GGACTCAGGGGTCTGCTCCGCCCAATGAGGCTGACGATGTCTCTAGTTGAACCTGGG 1440
DB 11380 GGACTCAGGGGTCTGCTCCGCCCAATGAGGCTGACGATGTCTCTAGTTGAACCTGGG 11439
QY 1441 AACCAATTTCCAAACTGTCCAGGGGCGGGGTTCGTCTCAATCTCCCTGGGCACTTTCA 1500
DB 11440 AACCAATTTCCAAACTGTGTCCAGGGGCGGGGTTCGTCTCAATCTCCCTGGGCACTTTCA 11499
QY 1501 TCTCTAAGCTCAGGGGCCCATCTCTTCTGTGACGCTGTGACCCAAATTTAGTCCGAAT 1560
DB 11500 TCTCTAAGCTCAGGGGCCCATCTCTTCTGTGACGCTGTGACCCAAATTTAGTCCGAAT 11559
QY 1561 AAATGAGAAG 1571
DB 11560 AAATGAGAAG 11570
```

RESULT 2
ADKS2482
ID ADKS2482 strand: DNA; 11570 BP.
XX
AC ADKS2482;
XX
AC
XX
DT 03-JUN-2004 (first entry)
XX
DE Human kallikrein 5 encoding sequence.
XX
KW kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.
XX
XX Homo sapiens.
OS
PN WO2004021008-A2.
XX
PD 11-MAR-2004.
XX
XX 28-AUG-2003; 2003WO-CA001310.
PF
XX 28-AUG-2002; 2002US-0407333P.
PR
XX (MOUN) MOUNT SINAI HOSPITAL.
XX (YOUS /) YOUSEF G.
PA

| Query Match | Similarity | Score | DB | Length | Matches | Conservative | Mismatches | Indels | Gaps | Other |
|-------------|------------|---|--------|--------|---------|--------------|------------|--------|------|-------|
| Best Local | 100.0% | 1571 | 12 | 11570 | 1571 | 0 | 0 | 0 | 0 | 0 |
| Matches | 1571 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Query | 1 | AGGAGGATTTGGGAAATTTGAAGACAGAAACACAAATTAATCCAGCGAATGGAATTTCTA | 60 | | | | | | | |
| Db | 10000 | AGGAGGATTTGGGAAATTTGAAGACAGAAACACAAATTAATCCAGCGAATGGAATTTCTA | 100559 | | | | | | | |
| Qy | 61 | TTGGAGATGATTTGCCCTTGAAGACATCGGCAATTCAGAGACATTTTGGTTGTCTA | 120 | | | | | | | |
| Db | 10060 | TTGGAGATGATTTGCCCTTGAAGACATCGGCAATTCAGAGACATTTTGGTTGTCTA | 101139 | | | | | | | |
| Qy | 121 | CAACTATATGAGAGGGGCAATTAATCGGCAATTAATGATGATGCGCAAGTGTCTTCAAC | 180 | | | | | | | |
| Db | 10120 | CAACTATATGAGAGGGGCAATTAATCGGCAATTAATGATGATGCGCAAGTGTCTTCAAC | 101799 | | | | | | | |
| Qy | 181 | ATGCTATGATGACACCGGACGGCTCCACAAACCAATTAATCAAGTTCAATGATGCCAC | 240 | | | | | | | |
| Db | 10180 | ATGCTATGATGACACCGGACGGCTCCACAAACCAATTAATCAAGTTCAATGATGCCAC | 102339 | | | | | | | |
| Qy | 241 | AGTGGCCAGATCGAGGAACCTCATCGAGGGGCTGGAACCGTATTTTGGCAAGGGAG | 300 | | | | | | | |
| Db | 10240 | AGTGGCCAGATCGAGGAACCTCATCGAGGGGCTGGAACCGTATTTTGGCAAGGGAG | 102939 | | | | | | | |
| Qy | 301 | GTAATAGATGGTGTGGTGAAGATTTGGGGAAGAAAGTGTGTCTCCAGTAAGAAATTA | 360 | | | | | | | |
| Db | 10300 | GTAATAGATGGTGTGGTGAAGATTTGGGGAAGAAAGTGTGTCTCCAGTAAGAAATTA | 103539 | | | | | | | |
| Qy | 361 | GGCCTGCACAGGCTGAGAGGGAGATGAGAGAAAGGAGCGGAGAGATACAGATGA | 420 | | | | | | | |
| Db | 10360 | GGCCTGCACAGGCTGAGAGGGAGATGAGAGAAAGGAGCGGAGAGATACAGATGA | 104139 | | | | | | | |
| Qy | 421 | GGGAGACAGGCTGGAACAGAAAGTAAGACGAATTCAGAGATGTGAGAGGAAGGCTCA | 480 | | | | | | | |
| Db | 10420 | GGGAGACAGGCTGGAACAGAAAGTAAGACGAATTCAGAGATGTGAGAGGAAGGCTCA | 104799 | | | | | | | |
| Qy | 481 | CAGACCCCCCGGAATGATGTGTGACAAACAGAAATCTGGAAGAGGAAGATGAGAGTGAG | 540 | | | | | | | |
| Db | 10480 | CAGACCCCCCGGAATGATGTGTGACAAACAGAAATCTGGAAGAGGAAGATGAGAGTGAG | 105339 | | | | | | | |
| Qy | 541 | AGTGAACAATGGGCTTAAGGTTGAACCTTGAGAGCGGCAATGATGTGCTACAGCTGTGA | 600 | | | | | | | |
| Db | 10540 | AGTGAACAATGGGCTTAAGGTTGAACCTTGAGAGCGGCAATGATGTGCTACAGCTGTGA | 105939 | | | | | | | |
| Qy | 601 | ATCCCAACACTTTGAGAGGCTGAGAGGTCGCAATTAATTTGAGGCGCAGAGATTCAGACCAAG | 660 | | | | | | | |
| Db | 10600 | ATCCCAACACTTTGAGAGGCTGAGAGGTCGCAATTAATTTGAGGCGCAGAGATTCAGACCAAG | 106539 | | | | | | | |

| | | | |
|----------|-------------|---|-------|
| QY | 661 | CTGGGCCAATAAGTGTAAACCCCGTCTCACAAAAAATATCAAAAAATTTACCGGGGT | 720 |
| Db | 10660 | CTGGCCCAATAGGTGAAACCCCGTCTCACAAAAAATATCAAAAAATTTAGCGGGGT | 10719 |
| QY | 721 | GGTATGGAACCTGTAGTCAACACTATTGGAGGCTGAGGCAAGAGATTGCTGAAC | 780 |
| Db | 10720 | GGTATGGAACCTGTATGTACAGCTATCTTGGAGGCTGAGGCAAGAGATTGCTGAAC | 10779 |
| QY | 781 | CCGGAGATGAGGCTGCAGTAGCTGAGGTCAAGGCCACTGGCTTCAACTGGGACA | 840 |
| Db | 10780 | CCGGAGATGAGGCTGCAGTAGCTGAGGTCAAGGCCACTGGCTTCAACTGGGACA | 10839 |
| QY | 841 | GAGTAAACCTCAATCTCAAAAAAAAAGCTGGAATTGAGAGTAAATTTAATAATT | 900 |
| Db | 10840 | GAGTAAACCTCAATCTCAAAAAAAAAGCTGGAATTGAGAGTAAATTTAATAATT | 10899 |
| QY | 901 | CTCCCTCTCTCTCTCTTTCCTGTCGTCTCATCTCTGTCCTTTTCTGCATTTCTTCATCT | 960 |
| Db | 10900 | CTCCCTCTCTCTCTCTTTCCTGTCGTCTCATCTCTGTCCTTTTCTGCATTTCTTCATCT | 10955 |
| QY | 961 | CTGTACTTTTCATCTCTGTGTGTCTGTCCATCTGCTTCTCATCTTATGGGATCTCTG | 1020 |
| Db | 10960 | CTGTACTTTTCATCTCTGTGTGTCTGTCCATCTGCTTCTCATCTTATGGGATCTCTG | 11019 |
| QY | 1021 | GGTCTCAATGTCCTTCCTGCCCCCATCTTGCCACATCTGTCGCTCTCATGCCCCCTT | 1080 |
| Db | 11020 | GGTCTCAATGTCCTTCCTGCCCCCATCTTGCCACATCTGTCGCTCTCATGCCCCCTT | 11079 |
| QY | 1081 | TCTCTCTGCAAGGATGATTCCTGGAGGAGCTGTGTGTCTGCATATGACTCCCTGACGGGACTC | 1140 |
| Db | 11080 | TCTCTCTGCAAGGATGATTCCTGGAGGAGCTGTGTGTCTGCATATGACTCCCTGACGGGACTC | 11139 |
| QY | 1141 | GTGTCTCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGGTCTTACACGAACCTTC | 1200 |
| Db | 11140 | GTGTCTCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGGTCTTACACGAACCTTC | 11199 |
| QY | 1201 | TGCAGTTTCAACCAAGTGAATCCAGAAAACATCCAGGCCACCTCTGAGTATCCACAGA | 1260 |
| Db | 11200 | TGCAGTTTCAACCAAGTGAATCCAGAAAACATCCAGGCCACCTCTGAGTATCCACAGA | 11255 |
| QY | 1261 | CTCAGCACACCGGATCTCCCACTGTGACAGGACAGCCCTGACATCTCCTTTCAAGCCT | 1320 |
| Db | 11260 | CTCAGCACACCGGATCTCCCACTGTGACAGGACAGCCCTGACATCTCCTTTCAAGCCT | 11319 |
| QY | 1321 | CATTCTTTCCCAAGATGTTGAAATGTTTCACTCTCCAGCCCCCTGACCCCAATGTCCTCT | 1380 |
| Db | 11320 | CATTCTTTCCCAAGATGTTGAAATGTTTCACTCTCCAGCCCCCTGACCCCAATGTCCTCT | 11379 |
| QY | 1381 | GGACTCAAGGATGTCCTTCCCCCACTTGGGCTGACCGTGTCTCTAGTTGAAACCTGTGG | 1440 |
| Db | 11380 | GGACTCAAGGATGTCCTTCCCCCACTTGGGCTGACCGTGTCTCTAGTTGAAACCTGTGG | 11439 |
| QY | 1441 | AACAATTTCAAAAATGTCACAGGAGCGAGGATGCTCAATCTCCCTGGGCACTTTCA | 1500 |
| Db | 11440 | AACAATTTCAAAAATGTCACAGGAGCGAGGATGCTCAATCTCCCTGGGCACTTTCA | 11499 |
| QY | 1501 | TCTCTCAAGCTCAGGGCCCATCTCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAAT | 1560 |
| Db | 11500 | TCTCTCAAGCTCAGGGCCCATCTCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAAT | 11555 |
| QY | 1561 | AAACTGAGAAG 1571 | |
| Db | 11560 | AAACTGAGAAG 11570 | |
| RESULT 3 | | | |
| ADR72623 | | | |
| ID | ADR72623 | standard; DNA; 11570 BP. | |
| XX | ADR72623; | | |
| DT | 02-DEC-2004 | (first entry) | |

DB 10000 AGGAGATATGGGAAATTGAGACAGAGAAACAAATATGATCCAAAGCAATGATTTCTTA 10059
QY 61 TTGGAGATGATTTGCCCCCTAGAGACAATGCAATACAGAGAGATTTTGGTTGTA 120
DB 10060 TTGGAGATGATTTGCCCCCTAGAGACAATGCAATACAGAGAGATTTTGGTTGTA 10119
QY 121 CAACATATATGAGAGGAGCATTTACTGCAACTATATGATATGATGCAAGTGTCTTCAAC 180
DB 10120 CAACATATATGAGAGGAGCATTTACTGCAACTATATGATATGATGCAAGTGTCTTCAAC 10179
QY 181 ATGCTATGATGCAACAGGAGGAGGCTTCAACAAACCATTTATCCAGTTGAGATGCCAC 240
DB 10180 ATGCTATGATGCAACAGGAGGAGGCTTCAACAAACCATTTATCCAGTTGAGATGCCAC 10239
QY 241 AGTGCCAGATGAGAGACCTCATCAAGGAGGCTGAGAACCGATTTTGGCAGAGAGGAG 300
DB 10240 AGTGCCAGATGAGAGACCTCATCAAGGAGGCTGAGAACCGATTTTGGCAGAGAGGAG 10299
QY 301 GTATAAGATGAGGATGAGAGATGAGAGAGAGAGAGAGATGTCAGTAAAGAAATTA 360
DB 10300 GTATAAGATGAGGATGAGAGATGAGAGAGAGAGAGAGATGTCAGTAAAGAAATTA 10359
QY 361 GGCCTGACAGGCTGAGAGGAGAGATGAGAGAGAGAGAGAGAGGCTGAGAGATACAGATGA 420
DB 10360 GGCCTGACAGGCTGAGAGGAGAGATGAGAGAGAGAGAGAGAGGCTGAGAGATACAGATGA 10419
QY 421 GGGAGACAGGCTGAGAGAGAGAGATGAGAGAGAGATGAGAGAGAGAGAGAGGCTGA 480
DB 10420 GGGAGACAGGCTGAGAGAGAGAGATGAGAGAGAGATGAGAGAGAGAGAGAGGCTGA 10479
QY 481 CAGACCCCCCGAAATGATGTGTGACAAACAGAAATCTGAGAGAGAGAGATGAGAGAG 540
DB 10480 CAGACCCCCCGAAATGATGTGTGACAAACAGAAATCTGAGAGAGAGAGATGAGAGAG 10539
QY 541 AGTGAACAATGGGCTTAAAGTTTGAACCTTGAAGGCCAGGATGTGGCTCACGCTGTA 600
DB 10540 AGTGAACAATGGGCTTAAAGTTTGAACCTTGAAGGCCAGGATGTGGCTCACGCTGTA 10599
QY 601 ATCCCAACCTTTGAGAGGCTGAGAGGAGGAGAAATCACTGAGGCAAGAGATGAGAGAG 660
DB 10600 ATCCCAACCTTTGAGAGGCTGAGAGGAGGAGAAATCACTGAGGCAAGAGATGAGAGAG 10659
QY 661 CCTGGCCAAATGATGAGAAACCCGCTCTTACAAAAAAATATACAAAAATTTAGCCGGGTGT 720
DB 10660 CCTGGCCAAATGATGAGAAACCCGCTCTTACAAAAAAATATACAAAAATTTAGCCGGGTGT 10719
QY 721 GGTGATGAGACCTGTATGTACAGCTACTTGGAGGCTGAGGAGAGAGAAATGCTTGAAC 780
DB 10720 GGTGATGAGACCTGTATGTACAGCTACTTGGAGGCTGAGGAGAGAGAAATGCTTGAAC 10779
QY 781 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCACTGCGCTCAACCTGGGAGACA 840
DB 10780 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCACTGCGCTCAACCTGGGAGACA 10839
QY 841 GAGTAAGACTCACTCAAAAAAAGCTGATTTGGAGTGAATATTAATTAACAT 900
DB 10840 GAGTAAGACTCACTCAAAAAAAGCTGATTTGGAGTGAATATTAATTAACAT 10899
QY 901 CTCCCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGATTTCTTCATCT 960
DB 10900 CTCCCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGATTTCTTCATCT 10959
QY 961 CTGTACTTTCCATCTCTGTGTCTGTCTCCATCTGCTCATGATGAGAGATCTG 1020
DB 10960 CTGTACTTTCCATCTCTGTGTCTGTCTCCATCTGCTCATGATGAGAGATCTG 11019
QY 1021 GGTCTCTCAATGCTCTCTTCTGCTCCCACTTGGCAGATCTGCTCTCAATGCCCTT 1080
DB 11020 GGTCTCTCAATGCTCTCTTCTGCTCCCACTTGGCAGATCTGCTCTCAATGCCCTT 11079
QY 1081 TCTCTCTGAGGATATTTCTGGGGGCTGCTGTGCAATGGCTTCCCTGACGAGATCT 1140

DB 11080 TCTCTCTGACAGGATGATTTGAGGAGGCTGTGTCTGTCAATGCTTCCCTGACGAGATC 11139
QY 1141 GTGTCCTGGGAGATTAACCTTGTGCTCCGAGCCCAACAGACCGGAGTGTCTACAGAACTC 1200
DB 11140 GTGTCCTGGGAGATTAACCTTGTGCTCCGAGCCCAACAGACCGGAGTGTCTACAGAACTC 11139
QY 1201 TGCAGATTCACCAATGATTCAGAGAAACCATTCAGGCCCACTCTGATCATTCAGAGA 1260
DB 11200 TGCAGATTCACCAATGATTCAGAGAAACCATTCAGGCCCACTCTGATCATTCAGAGA 11259
QY 1261 CTCAGACACCGGACATCCCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 11260 CTCAGACACCGGACATCCCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11319
QY 1321 CATTCCTTCCAGAGATGTTGAATGTTCACTCTCTCAGAGCCCTGAGCCCACTGTCTCT 1380
DB 11320 CATTCCTTCCAGAGATGTTGAATGTTCACTCTCTCAGAGCCCTGAGCCCACTGTCTCT 11379
QY 1381 GGACTCAGGAGTGTCTTCCCAATGAGGCTGACCGTGTCTCTTATGTAACCTGGG 1440
DB 11380 GGACTCAGGAGTGTCTTCCCAATGAGGCTGACCGTGTCTCTTATGTAACCTGGG 11439
QY 1441 AACATTTCCAAACTGTTCAGAGGAGGAGGAGTGTCTCTCAATCTCTCTGGGCACTTTCA 1500
DB 11440 AACATTTCCAAACTGTTCAGAGGAGGAGGAGTGTCTCTCAATCTCTCTGGGCACTTTCA 11499
QY 1501 TCCCTCAGGCTCAGGAGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAAT 1560
DB 11500 TCCCTCAGGCTCAGGAGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAAT 11559
QY 1561 AAACCTGAGAG 1571
DB 11560 AAACCTGAGAG 11570

RESULT 5
ACH74409
ID ACH74409 standard; DNA; 567 BP.
XX
AC ACH74409;
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #7604.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
OS Homo sapiens.
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
PI Penn SG, Rank DR, Hanzel DK;
PI
DR WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 15; SEQ ID NO 7604; 80bp; English.
XX

CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subsequence, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704

CC XX Sequence 567 BP; 105 A; 197 C; 125 G; 140 T; 0 U; 0 Other;

CC Query Match 34.4%; Score 541; DB 12; Length 567;
CC Best Local Similarity 100.0%; Pred. No. 7e-128;

CC Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1031 GTGCTCTTGGCCCACTTTGGACATCTGCTGCAATGAGCCCTTCTCTCTGCG 1090
CC 1 GTCTCTTCTGCCCCACCTTGGACATCTGCTCTCTCTGACCCCTTCTCTCTGCG 60
CC 1091 AGGGTATTCCTGGGGGCTGTGTCTGCAATGAGCTCTCTGACAGGACTGCTGTCTGGG 1150
CC 61 AGGGTATTCCTGGGGGCTGTGTCTGCAATGAGCTCTCTGACAGGACTGCTGTCTGGG 120
CC 1151 GAGATTACCTTGTGCCCCGCGCAACAGACCGGGTGTCTACAGCACTCTGCAAGTTCA 1210
CC 121 GAGATTACCTTGTGCCCCGCGCAACAGACCGGGTGTCTACAGCACTCTGCAAGTTCA 180
CC 1211 CCAAGTGAATCAGAGAAACATCCAGGCAACTCTGATGATCCAGAGACTGACAGAC 1270
CC 181 CCAAGTGAATCAGAGAAACATCCAGGCAACTCTGATGATCCAGAGACTGACAGAC 240
CC 1271 CGGATCCCACTGTCTGACAGGCAAGCCCTGACACTCTTTCAGACCTTATCTCTTC 1330
CC 241 CGGATCCCACTGTCTGACAGGCAAGCCCTGACACTCTTTCAGACCTTATCTCTTC 300
CC 1331 CAGAGATGTTGAGAAATTTATCTCTCAGCCCTGACCCCATGCTCTCTGAGACTCAGGG 1390
CC 301 CAGAGATGTTGAGAAATTTATCTCTCAGCCCTGACCCCATGCTCTCTGAGACTCAGGG 360
CC 1391 TCTGCTTCCCCCACTTGGGCTGACCGGTCTCTCTAGTTGAAACCTGGGAAACAATTTC 1450
CC 361 TCTGCTTCCCCCACTTGGGCTGACCGGTCTCTCTAGTTGAAACCTGGGAAACAATTTC 420
CC 1451 AAAAATGTCAGAGGCGGGGGTGTGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGCT 1510
CC 421 AAAAATGTCAGAGGCGGGGGTGTGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGCT 480

CC QY 1511 CAGGGCCCATCTCTTCTGACAGCTGACCCAAATTTAGTCCAGAAATTAACAGAA 1570
CC DB 481 CAGGGCCCATCTCTTCTGACAGCTGACCCAAATTTAGTCCAGAAATTAACAGAA 540
CC QY 1571 G 1571
CC DB 541 G 541
CC
CC RESULT 6
CC ABL63579/c
CC ID ABL63579 standard; DNA; 586 BP.
CC AC ABL63579;
CC XX
CC DT 15-MAY-2002 (first entry)
CC XX
CC DE Breast cancer related gene sequence SEQ ID NO:1916.
CC KM Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
CC KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
CC KM cytotoxic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
CC KM gene; ds.
CC XX
CC OS Homo sapiens.
CC XX
CC PN W0200194629-A2.
CC PD 13-DEC-2001.
CC XX
CC PF 30-MAY-2001; 2001NC-US010838.
CC XX
CC PR 05-JUN-2000; 2000US-0209473P.
CC PR 05-JUN-2000; 2000US-0209531P.
CC PR 18-SEP-2000; 2000US-023133P.
CC PR 18-SEP-2000; 2000US-0233617P.
CC PR 20-SEP-2000; 2000US-0234009P.
CC PR 20-SEP-2000; 2000US-0234034P.
CC PR 20-SEP-2000; 2000US-0234052P.
CC PR 22-SEP-2000; 2000US-0234509P.
CC PR 22-SEP-2000; 2000US-0234567P.
CC PR 25-SEP-2000; 2000US-0234923P.
CC PR 25-SEP-2000; 2000US-0234924P.
CC PR 25-SEP-2000; 2000US-0235077P.
CC PR 25-SEP-2000; 2000US-0235082P.
CC PR 25-SEP-2000; 2000US-0235134P.
CC PR 25-SEP-2000; 2000US-0235280P.
CC PR 26-SEP-2000; 2000US-0235637P.
CC PR 26-SEP-2000; 2000US-0235638P.
CC PR 27-SEP-2000; 2000US-0235711P.
CC PR 27-SEP-2000; 2000US-0235720P.
CC PR 27-SEP-2000; 2000US-0235840P.
CC PR 27-SEP-2000; 2000US-0235863P.
CC PR 28-SEP-2000; 2000US-0236028P.
CC PR 28-SEP-2000; 2000US-0236032P.
CC PR 28-SEP-2000; 2000US-0236033P.
CC PR 28-SEP-2000; 2000US-0236034P.
CC PR 28-SEP-2000; 2000US-0236109P.
CC PR 28-SEP-2000; 2000US-0236111P.
CC PR 29-SEP-2000; 2000US-0236642P.
CC PR 29-SEP-2000; 2000US-0236681P.
CC PR 29-SEP-2000; 2000US-0236891P.
CC PR 02-OCT-2000; 2000US-0237172P.
CC PR 02-OCT-2000; 2000US-0237173P.
CC PR 02-OCT-2000; 2000US-0237278P.
CC PR 02-OCT-2000; 2000US-0237294P.
CC PR 02-OCT-2000; 2000US-0237295P.
CC PR 02-OCT-2000; 2000US-0237316P.
CC PR 02-OCT-2000; 2000US-0237425P.
CC PR 03-OCT-2000; 2000US-0237598P.
CC PR 03-OCT-2000; 2000US-0237604P.
CC PR 03-OCT-2000; 2000US-0237606P.
CC PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-024867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 1916; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I1) of a signature gene set, where (I1)
 CC comprises a sequence (S1) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S1), where a change in
 CC expression is indicative of anti-neoplastic activity. (I1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour
 XX
 SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;
 Query Match 30.7%; Score 482; DB 6; Length 586;
 Best Local Similarity 100.0%; Pred. No. 8.3e-113;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1090 CAGGGGATTTCTGGGGGCTGTGTCGAAAGCTCCCTGAGGAGCTCGTCTCGG 1149
 DB 501 CAGGGGATTTCTGGGGGCTGTGTCGAAAGCTCCCTGAGGAGCTCGTCTCGG 442
 QY 1150 GGAGATTACCTTTGCTCCCGGCCAACAGACGGGGTGTCTACACGAACCTTCAGATTG 1209
 DB 441 GGAGATTACCTTTGCTCCCGGCCAACAGACGGGGTGTCTACACGAACCTTCAGATTG 382
 QY 1210 ACCAAGTGAATCCAGAAACATTCAGGCCCACTCCTGAGTCATCCAGAGCTCAGCACA 1269
 DB 381 ACCAAGTGAATCCAGAAACATTCAGGCCCACTCCTGAGTCATCCAGAGCTCAGCACA 322
 QY 1270 CCGGCATCCCACTCTGCTGAGGAGAGCCCTGACACTCTTTCAACCTCTATTCTTC 1329
 DB 321 CCGGCATCCCACTCTGCTGAGGAGAGCCCTGACACTCTTTCAACCTCTATTCTTC 262
 QY 1330 CCAGAGATTTGGAATGTTCACTCTCCAGCCCTGAGCCCACTGCTCTGAGACTCAGG 1389
 DB 261 CCAGAGATTTGGAATGTTCACTCTCCAGCCCTGAGCCCACTGCTCTGAGACTCAGG 202
 QY 1390 GTCTGTCCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTTGGGAACAATTC 1449
 DB 201 GTCTGTCCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTTGGGAACAATTC 142
 QY 1450 CAAAATGTCAGGGGGGGGGTGGCTTCAATCTCCCTGGGCACTTTTCACTCTCAAGC 1509
 DB 141 CAAAATGTCAGGGGGGGGGTGGCTTCAATCTCCCTGGGCACTTTTCACTCTCAAGC 82
 QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
 DB 81 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 22
 QY 1570 AG 1571

DB 21 AG 20
 RESULT 7
 ABL63996/C
 ID ABL63996 standard; DNA; 586 BP.
 XX
 AC ABL63996;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Breast cancer related gene sequence SEQ ID NO:2333.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
 KW gene; db.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0231133P.
 PR 18-SEP-2000; 2000US-023617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, 'Weaver Z;
XX
DR WPI; 2002-188264/24.
...

PT Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

PS Claim 1; SEQ ID NO 2333; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABIL61664 to ABIL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytostatic activity and can be used in gene therapy. M1 can be used for screening anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour

SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;

| | | | | |
|-------------|-------|------------|-------|-------------|
| Query Match | 30.7% | Score 482; | DB 6; | Length 586; |
|-------------|-------|------------|-------|-------------|

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| Qy | 1090 | CAGGGTATTTCTGGGGGGGCGTGTGGTCTGCAGATGGCTCCCTGCAGAGGAGCTGATGCTCGG | 1149 |
| Db | 501 | CAGGGTATTTCTGGGGGGGCGTGTGGTCTGCAGATGGCTCCCTGCAGAGGAGCTGATGCTCGG | 442 |
| Qy | 1150 | GGAGATTACCTTGTGTGCCCGGGCCACAGACCGGGGTGTCTACAGAACCTTCTGCAGATTG | 1209 |
| Db | 441 | GGAGATTACCTTGTGTGCCCGGGCCACAGACCGGGGTGTCTACAGAACCTTCTGCAGATTG | 382 |
| Qy | 1210 | ACCAAGTGGATCCAGAGAAACCATCAGGGCCAACTCCTGAGCATCCCGAGACTCAGACACA | 1269 |
| Db | 381 | ACCAAGTGGATCCAGAGAAACCATCAGGGCCAACTCCTGAGCATCCCGAGACTCAGACACA | 322 |
| Qy | 1270 | CCGGCATCCCCACCTGCTGCAGGAGCAGCCCTGACATCTCTTTGAGACCTTCATTTCCTTC | 1329 |
| Db | 321 | CCGGCATCCCCACCTGCTGCAGGAGCAGCCCTGACATCTCTTTGAGACCTTCATTTCCTTC | 262 |
| Qy | 1330 | CCAGAGATGTTGAGAAATGTTTATCTCTCCAGCCCCCTGACCCCAATGTCTCTCTGAGACTCAGG | 1389 |
| Db | 261 | CCAGAGATGTTGAGAAATGTTTATCTCTCCAGCCCCCTGACCCCAATGTCTCTCTGAGACTCAGG | 202 |
| Qy | 1390 | GTCGTCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCTGGGAAACAATTC | 1449 |
| Db | 201 | GTCGTCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCTGGGAAACAATTC | 142 |
| Qy | 1450 | CAAAACTGTCCAGGGCGGGGGGTTGCGTCTCAATCTCCTGGGGACATTTCACTCTCAAGC | 1509 |
| Db | 141 | CAAAACTGTCCAGGGCGGGGGGTTGCGTCTCAATCTCCTGGGGACATTTCACTCTCAAGC | 82 |
| Qy | 1510 | TCAGGGGCCATCCCTTCTCTGACAGCTGAAACCAATTAGTCCAGAAATTAACCTGAGA | 1569 |
| Db | 81 | TCAGGGGCCATCCCTTCTCTGACAGCTGAAACCAATTAGTCCAGAAATTAACCTGAGA | 22 |
| Qy | 1570 | AG 1571 | |
| Db | 21 | AG 20 | |

RESULT 8
AAC79469/c

| ID | AAC79469 standard; cDNA; 735 BP |
|----|---------------------------------|
| XX | |
| AC | AAC79469; |

DT 07-FEB-2001 (first entry)

DE cDNA sequence of human breast tumour clone B541S

KW Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;

25

[illegible]

XX
3
3
3
3

XX 10-APR-2000
DE

XX 00 - APB - 1000 - 0011C - 00388650
DB

XX PR 02-JUL-1999; 99US-00346327. XX

PA (CORI-) CORIXA CORP.
XX

P1 Reed SG, Xu J, Dillon DC,
XX

DK WP1; 2000-638568/61.
XX

PT breast cancer protein useful in the detection and treatment of breast

F1 cancer, XX

XX
ES
XX

CLAIM #; Page 00; English

CC The present sequence was isolated from a breast tumour cDNA library. It
CC is provided in a specification relating to compounds for immunotherapy
CC and diagnosis of breast cancer. Breast tumour antigens and the
CC polynucleotides that encode them may be used in the production of a
CC pharmaceutical composition to be used in the treatment of breast cancer
CC proliferated T cells and incubated antigen presenting cells are also
CC required. The polypeptides and polynucleotides may also be used to
CC produce a vaccine

SQ Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;

| | | | | |
|-------------|-------|------------|-------|-------------|
| Query Match | 30.7% | Score 482; | DB 3; | Length 735; |
|-------------|-------|------------|-------|-------------|

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| Qy | 1090 | CAGGGTATTTCTGGGGGGGGCGTGGTGTGCAATGGCTCCCTGCAAGGGAGCTGTGTCTCGG | 1145 |
| Db | 500 | CAGGGTATTTCTGGGGGGGGCGTGGTGTGCAATGGCTCCCTGCAAGGGAGCTGTGTCTCGG | 441 |
| Qy | 1150 | GGAGATTACCTTGTGTGCCCCGGGCCCAAGAGCCGGGTGTCTACAGGAACCTTGTGCAATTTC | 1205 |
| Db | 440 | GGAGATTACCTTGTGTGCCCCGGGCCCAAGAGCCGGGTGTCTACAGGAACCTTGTGCAATTTC | 381 |
| Qy | 1210 | ACCAAGTGGATTCAGAGAAACCATCAGGSCCAATCTGTAGTCATTCGCCAGACTGACACA | 1265 |
| Db | 380 | ACCAAGTGGATTCAGAGAAACCATCAGGSCCAATCTGTAGTCATTCGCCAGACTGACACA | 321 |
| Qy | 1270 | CCGGCATTCCCACTTGCTGCAAGGAGCAGCCCTGACATCTCTTTCAAGACCTTCATTTCTTTC | 1322 |
| Db | 320 | CCGGCATTCCCACTTGCTGCAAGGAGCAGCCCTGACATCTCTTTCAAGACCTTCATTTCTTTC | 261 |
| Qy | 1330 | CCAAGATGTTGAGAAATGTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG | 1388 |
| Db | 260 | CCAAGATGTTGAGAAATGTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG | 201 |
| Qy | 1390 | GTCGTCTTCCCCCAATTGGGCTGACCGTGTCTCTTAGTTGAACCCCTGGGAACAATTC | 1448 |
| Db | 200 | GTCGTCTTCCCCCAATTGGGCTGACCGTGTCTCTTAGTTGAACCCCTGGGAACAATTC | 141 |

QY 1450 CAAAACTGTCAGAGGCGGGGTTGCCTCAATCTCCCTGAGGCACTTTCATCTCAAGC 1509
DB 140 CAAAATGTCAGAGGCGGGGTTGCCTCAATCTCCCTGAGGCACTTTCATCTCAAGC 81
QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTTAGTCCGAAATAAAGTGA 1569
DB 80 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTTAGTCCGAAATAAAGTGA 21
QY 1570 AG 1571
DB 20 AG 19
RESULT 9
ID ABR29013/c
ABK29013 standard; cDNA; 735 BP.
XX ABR29013;
DT 23-APR-2002 (first entry)
XX Human breast tumour polypeptide cDNA clone #42.
DE Human breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW Immunostimulant.
XX Homo sapiens.
OS
PN MO200198339-A2.
XX
PD 27-DEC-2001.
PF 12-JUN-2001; 2001WO-US019032.
XX
PR 22-JUN-2000; 2000US-00602877.
PR 12-OCT-2000; 2000US-00687507.
PR 06-FEB-2001; 2001US-00778381.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
XX WPI; 2002-147792/19.
DR
XX Polynucleotides encoding breast polypeptides, useful for treating
PT breast cancer or stimulating an immune response.
XX
PS Claim 1; Page 140; 150pp; English.
XX
CC The invention relates to polynucleotides encoding breast tumour
CC polypeptides. The sequences are useful for treating cancer, preferably
CC breast cancer, in a patient or for stimulating an immune response. The
CC polynucleotides and polypeptides are also useful in the diagnosis and
CC monitoring of breast cancer. A method for detecting the presence of a
CC cancer in a patient, comprises obtaining a biological sample from the
CC patient, contacting the biological sample with a binding agent that binds
CC to a breast tumour polypeptide, detecting in the sample an amount of
CC polypeptide that binds to the binding agent, and comparing the amount of
CC polypeptide to a predetermined cut-off value, therefore determining the
CC presence of a cancer in the patient. Sequences ABR28920-ABK29025
CC represent cDNA clones encoding human breast tumour polypeptides of the
CC invention
XX
SQ Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 9e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGGATTCCTGGGGGCGCTGGTCTCAATGCTCCCTGAGGAGCTCGTCTCGG 1149
DB 500 CAGGGGATTCCTGGGGGCGCTGGTCTCAATGCTCCCTGAGGAGCTCGTCTCGG 441

QY 1150 GGAGATTACCTTGTGCCCCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTGTGCCCCCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 381
QY 1210 ACCAAGTGAATCAGAGAAACCTTCCAGGCTCAATCTCTGAGTATCCAGAGACTGACACA 1269
DB 380 ACCAAGTGAATCAGAGAAACCTTCCAGGCTCAATCTCTGAGTATCCAGAGACTGACACA 321
QY 1270 CCGGATCCCGACCTGCTGACAGGAGACAGCCCTGACACTCTTTGACACCTCATTCCTTC 1329
DB 320 CCGGATCCCGACCTGCTGACAGGAGACAGCCCTGACACTCTTTGACACCTCATTCCTTC 261
QY 1330 CCAGAGATGTTGAGATGTTCACTCTCTCAGCCCTGACACCCCATGCTCTCTGAGACTGAG 1389
DB 260 CCAGAGATGTTGAGATGTTCACTCTCTCAGCCCTGACACCCCATGCTCTCTGAGACTGAG 201
QY 1390 GTCTGCTTCCCCCAATTTGGGTGACCGTGTCTCTAGTTGAACCCCTGGAGAACTTTC 1449
DB 200 GTCTGCTTCCCCCAATTTGGGTGACCGTGTCTCTAGTTGAACCCCTGGAGAACTTTC 141
QY 1450 CAAACTGTCAGAGGCGGGGTTGCGTCAATCTCCCTGAGGCACTTTCATCTCAAGC 1509
DB 140 CAAACTGTCAGAGGCGGGGTTGCGTCAATCTCCCTGAGGCACTTTCATCTCAAGC 81
QY 1510 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTTAGTCCGAAATAAAGTGA 1569
DB 80 TCAGGGCCCATCCCTTCTCTGCAAGCTCTGACCCAAATTTAGTCCGAAATAAAGTGA 21
QY 1570 AG 1571
DB 20 AG 19

RESULT 10
ID AAX84240/c
AAX84240 standard; cDNA; 738 BP.
XX
AC AAX84240;
XX
DT 08-SEP-1999 (first entry)
XX
DE DNA encoding human breast tumour protein immunogenic fragment.
XX
KW Breast tumour protein; immunogenic fragment; vaccine; detection;
KW breast cancer development; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9933869-A2.
XX
PD 08-JUL-1999.
PP 22-DEC-1998; 98WO-US027416.
XX
PR 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J;
XX
DR WPI; 1999-405486/34.
XX
PT New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
PS Claim 3; Page 70; 70pp; English.
XX
CC This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
CC useful in vaccines and pharmaceutical compositions for manufacture of

CC medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 738 BP; 161 A; 179 C; 226 G; 172 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 2; Length 738;
Best Local Similarity 100.0%; Pred. No. 9.1e-113; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGACAGGACTGTCCTGG 1149
DB 503 CAGGGTATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGACAGGACTGTCCTGG 444
QY 1150 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACACGAACTTCTGCAATTC 1209
DB 443 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACACGAACTTCTGCAATTC 384
QY 1210 ACCAAGTGCATCCAGGAAACCATCCAGGCCCACTCTGAGTCACTCCAGACTCAGACA 1269
DB 383 ACCAAGTGCATCCAGGAAACCATCCAGGCCCACTCTGAGTCACTCCAGACTCAGACA 324
QY 1270 CCGGCATCCCGACCTGTGTCGAGGAGACCCCTGACATCCCTTTCAGACCCCTCATTCCTTC 1329
DB 323 CCGGCATCCCGACCTGTGTCGAGGAGACCCCTGACATCCCTTTCAGACCCCTCATTCCTTC 264
QY 1330 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1389
DB 263 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 204
QY 1390 GTGTGCTTCCCGCCCATTTGGGCTGACCGTGTCTCTAGTGAACCTTGGGAACAATTTTC 1449
DB 203 GTGTGCTTCCCGCCCATTTGGGCTGACCGTGTCTCTAGTGAACCTTGGGAACAATTTTC 144
QY 1450 CAAAAGTGTCCAGGGGCGGGGTTGCGTCAATCTCCCTGGGAGCATTTTCACTCTCAGC 1509
DB 143 CAAAAGTGTCCAGGGGCGGGGTTGCGTCAATCTCCCTGGGAGCATTTTCACTCTCAGC 84
QY 1510 TCAAGGCCCATCCCTTCTCTGACAGCTGACCCAAATTTAGTCCAGAAATMAACTGAGA 1569
DB 83 TCAAGGCCCATCCCTTCTCTGACAGCTGACCCAAATTTAGTCCAGAAATMAACTGAGA 24
QY 1570 AG 1571
DB 23 AG 22
RESULT 11
ADB80489
ID ADB80489 standard; DNA; 1260 BP.
XX
XX ADB80489;
AC
XX
DT 04-DEC-2003 (first entry)
XX
XX Ovarian cancer-associated transcript #27.
DE
XX
XX Cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 43..924
FT CDS /*tag= a

XX
PN WO2002102235-A2.
XX
XX 27-DEC-2002.
PF
XX 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC;
PI
XX
XX WPI; 2003-167431/16.
DR P-PSDB; ADB80490.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS Claim 10; Page 292; 332pp; English.
XX
CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 10; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.1e-112; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGACAGGACTGTCCTGG 1149
DB 766 CAGGGTATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGACAGGACTGTCCTGG 825
QY 1150 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACACGAACTTCTGCAATTC 1209
DB 826 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACACGAACTTCTGCAATTC 885
QY 1210 ACCAAGTGCATCCAGGAAACCATCCAGGCCCACTCTGAGTCACTCCAGACTCAGACA 1269
DB 886 ACCAAGTGCATCCAGGAAACCATCCAGGCCCACTCTGAGTCACTCCAGACTCAGACA 945
QY 1270 CCGGCATCCCGACCTGTGTCGAGGAGACCCCTGACATCCCTTTCAGACCCCTCATTCCTTC 1329
DB 946 CCGGCATCCCGACCTGTGTCGAGGAGACCCCTGACATCCCTTTCAGACCCCTCATTCCTTC 1005
QY 1330 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1389
DB 1006 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1065
QY 1390 GTGTGCTTCCCGCCCATTTGGGCTGACCGTGTCTCTAGTGAACCTTGGGAACAATTTTC 1449
DB 1066 GTGTGCTTCCCGCCCATTTGGGCTGACCGTGTCTCTAGTGAACCTTGGGAACAATTTTC 1125
QY 1450 CAAAAGTGTCCAGGGGCGGGGTTGCGTCAATCTCCCTGGGAGCATTTTCACTCTCAGC 1509

Db 1126 CAAAGTGTCCAGGCGGGGTTGCGTTCATCTCCCTGGGCGACCTTCATCTCAAGC 1185
 QY 1510 TCAGGGCCCATCCCTTCTGTGAGCTGTGACCAATTTAGTCCGAATAAATGAGG 1569
 Db 1186 TCAGGGCCCATCCCTTCTGTGAGCTGTGACCAATTTAGTCCGAATAAATGAGG 1245
 QY 1570 AG 1571
 ||
 1246 AG 1247
 RESULT 12
 ADN39197
 ID ADN39197 standard; cDNA; 1260 BP.
 XX
 AC ADN39197;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.
 XX
 KM Human; differential expression; cancer; angiogenic disorder;
 KM fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KM inflammatory disease; autoimmune disease;
 KM retinal neovascularisation syndrome; scarring; uterine fibroid;
 KM detection; diagnosis; prognosis; drug screening; drug targeting;
 KM wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KM vulnary; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368099P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR P-PSDB; ADN39198.
 XX
 WIPI; 2003-468649/44.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO 515; 1385bp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;
 Query Match 30.7%; Score 482; DB 11; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 1.1e-112;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1090 CAGGGTATTCGGGGGCGCTGTGTCATGCTCCCTGCAAGGACTCGTGTCCGG 1149
 |||||
 Db 766 CAGGGTATTCGGGGGCGCTGTGTCATGCTCCCTGCAAGGACTCGTGTCCGG 825
 QY 1150 GGAGATTACCTTGTGCGCGGCCCAAGACCGGGTGTCTACAGCAACTCTGCAAGTTC 1209
 |||||
 Db 826 GGAGATTACCTTGTGCGCGGCCCAAGACCGGGTGTCTACAGCAACTCTGCAAGTTC 885
 QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTGATGATCCAGGACTCGACACA 1269
 |||||
 Db 886 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTGATGATCCAGGACTCGACACA 945
 QY 1270 CCGGATCCCGACCTGTGCAAGGACAGCCCTGACACTCTTGAGACCCGATTCCTTC 1329
 |||||
 Db 946 CCGGATCCCGACCTGTGCAAGGACAGCCCTGACACTCTTGAGACCCGATTCCTTC 1005
 QY 1330 CCAGAGATGTTGAGATGTTCTCTCCAGGCGCCCTGACCCGATGCTCTGAGACTCAG 1389
 |||||
 Db 1006 CCAGAGATGTTGAGATGTTCTCTCCAGGCGCCCTGACCCGATGCTCTGAGACTCAG 1065
 QY 1390 GTCTGCTTCCCGCAATTTGGGCTGACCGTGTCTCTGATGTAACCTTGGGAACAATTTC 1449
 |||||
 Db 1066 GTCTGCTTCCCGCAATTTGGGCTGACCGTGTCTCTGATGTAACCTTGGGAACAATTTC 1125
 QY 1450 CAAAGTGTCCAGGCGGGGCTGTGCTCATCTCCCGGGGCACTTTCATCTCAAGC 1509
 |||||
 Db 1126 CAAAGTGTCCAGGCGGGGCTGTGCTCATCTCCCGGGGCACTTTCATCTCAAGC 1185
 QY 1510 TCAGGGCCCATCCCTTCTGTGAGCTGTGACCAATTTAGTCCGAATAAATGAGG 1569
 |||||
 Db 1186 TCAGGGCCCATCCCTTCTGTGAGCTGTGACCAATTTAGTCCGAATAAATGAGG 1245
 QY 1570 AG 1571
 ||
 Db 1246 AG 1247
 RESULT 13
 ADR25550
 ID ADR25550 standard; DNA; 1260 BP.
 XX
 AC ADR25550;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Breast cancer prognosis marker #1411.
 XX
 KM de; breast cancer; prognosis; gene expression; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN WO2004065545-A2.

QY 1330 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCATGTCCTCGACTCAGG 1389
|||
Db 1051 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCATGTCCTCGACTCAGG 1110
QY 1390 GTCGTCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAACAAATTC 1449
|||
Db 1111 GTCGTCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAACAAATTC 1170
QY 1450 CAAACATGTCAGAGGGGGGCTTGCCTCCTCAATCTCCCTGGGACCTTTCATCTCAAGC 1509
|||
Db 1171 CAAACATGTCAGAGGGGGGCTTGCCTCCTCAATCTCCCTGGGACCTTTCATCTCAAGC 1230
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
|||
Db 1231 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1290
QY 1570 AG 1571
||
Db 1291 AG 1292

RESULT 15

AAK57989 standard; DNA; 1302 BP.

AAK57989;

19-JUL-1999 (first entry)

Human BS247 specific polynucleotide #7.

BS247; detection; diagnosis; breast cancer; atypical hyperplasia;
fibroadenoma; cystic breast disease; gene therapy; ss.

Homo sapiens.

MO9922027-A1.

06-MAY-1999.

28-OCT-1998; 98MO-US022906.

28-OCT-1997; 97US-0063431P.

28-OCT-1997; 97US-00968838.

(ABBO) ABBOT LAB.

Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC,
Stroupe SD, Yu H;

WPI, 1999-312977/26.

Breast tissue derived cDNA contig and consensus polypeptide sequence.

Claim 1; Page 105; 112pp; English.

This sequence is a BS247 specific polynucleotide. The invention relates to a method of detecting the presence of a target BS247 polynucleotide, especially mRNA, in a test sample. BS247 polynucleotides are derived from breast tissue. The polynucleotides, polypeptides or antibodies are useful for providing information leading to the detection, diagnosis, staging, monitoring, prognosis, in vivo imaging, prevention or treatment, determining predisposition to, diseases and conditions of the breast, such as breast cancer, atypical hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or gene therapy for breast cancer, can be based on these identified gene sequences and the efficacy of any particular therapy can be monitored. The BS247-derived reagents are advantageous for detection of breast cancer due to their specificity. The reagents also provide an alternative, non-surgical diagnostic method capable of detecting early stage breast disease, such as cancer

SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 2; Length 1302;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATGATTTGGGGGGGCTGTGTGTCGCAATGCTCCCTGACAGGACTCGTGTCTTG 1149
|||
Db 811 CAGGATGATTTGGGGGGGCTGTGTGTCGCAATGCTCCCTGACAGGACTCGTGTCTTG 870
QY 1150 GAGATTTACCTTTGTGCGCGCCCAACAGACCGGGTGTACAGAACTCTGCAAGTTC 1209
|||
Db 871 GAGATTTACCTTTGTGCGCGCCCAACAGACCGGGTGTACAGAACTCTGCAAGTTC 930
QY 1210 ACCAATGATTCAGAGAAACCATCCAGGCAACCTCTGATGATCCAGAGACTCAGACA 1269
|||
Db 931 ACCAATGATTCAGAGAAACCATCCAGGCAACCTCTGATGATCCAGAGACTCAGACA 990
QY 1270 CCGGATCCCACTGCTGACAGGGAAGCCCTGACATCTCTTTCAGACCTTCATTCCTTC 1329
|||
Db 991 CCGGATCCCACTGCTGACAGGGAAGCCCTGACATCTCTTTCAGACCTTCATTCCTTC 1050
QY 1330 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCATGTCCTCGACTCAGG 1389
|||
Db 1051 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCATGTCCTCGACTCAGG 1110
QY 1390 GTCGTCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAACAAATTC 1449
|||
Db 1111 GTCGTCTTCCCAACATTTGGGCTGACCGGTCTCTCTAGTTGAACCTTGGAACAAATTC 1170
QY 1450 CAAACATGTCAGAGGGGGGCTTGCCTCCTCAATCTCCCTGGGACCTTTCATCTCAAGC 1509
|||
Db 1171 CAAACATGTCAGAGGGGGGCTTGCCTCCTCAATCTCCCTGGGACCTTTCATCTCAAGC 1230
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
|||
Db 1231 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1290
QY 1570 AG 1571
||
Db 1291 AG 1292

Search completed: February 25, 2005, 20:15:07

Job time : 802.397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 18:05:22 : Search time 6692.88 Seconds
(without alignments)
11373.754 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 aaggagagctcgggaactga.....ccagaataactgagaag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbank:.*
1: gb_ba:.*
2: gb_hg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sta:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1 | 1571 | 100.0 | 11570 | 6 | CQ788219 | CQ788219 Sequence |
| 2 | 1571 | 100.0 | 11570 | 6 | CQ874885 | CQ874885 Sequence |
| 3 | 1571 | 100.0 | 11570 | 6 | CQ874961 | CQ874961 Sequence |
| 4 | 1571 | 100.0 | 11570 | 9 | AF135028 | AF135028 Homo sapi |
| 5 | 1571 | 100.0 | 107487 | 9 | AF011483 | AF011483 Homo sapi |
| 6 | 1571 | 100.0 | 230000 | 9 | AF243527 | AF243527 Homo sapi |
| 7 | 1555.2 | 99.0 | 217346 | 2 | AC027602 | AC027602 Homo sapi |
| 8 | 1482 | 94.3 | 200792 | 2 | AC130782 | AC130782 Pan trogl |
| 9 | 482 | 30.7 | 586 | 6 | AX331407 | AX331407 Sequence |
| 10 | 482 | 30.7 | 586 | 6 | AX331824 | AX331824 Sequence |
| 11 | 482 | 30.7 | 586 | 11 | G41906 | G41906 SHC-56840 |
| 12 | 482 | 30.7 | 735 | 6 | AX429955 | AX429955 Sequence |
| 13 | 482 | 30.7 | 738 | 6 | BD139877 | BD139877 Compound |
| 14 | 482 | 30.7 | 1143 | 6 | CQ720600 | CQ720600 Sequence |
| 15 | 482 | 30.7 | 1301 | 9 | AY461805 | AY461805 Homo sapi |
| 16 | 482 | 30.7 | 1370 | 9 | AY279380 | AY279380 Homo sapi |
| 17 | 482 | 30.7 | 1381 | 6 | BD107879 | BD107879 36 human |
| 18 | 482 | 30.7 | 1387 | 6 | CQ874884 | CQ874884 Sequence |
| 19 | 482 | 30.7 | 1387 | 6 | CQ874960 | CQ874960 Sequence |

| | | | | | | |
|----|-------|------|--------|----|----------|--------------------|
| 20 | 482 | 30.7 | 1387 | 9 | AF168768 | AF168768 Homo sapi |
| 21 | 482 | 30.7 | 1438 | 6 | AY279381 | AY279381 Homo sapi |
| 22 | 482 | 30.7 | 1439 | 6 | BD107880 | BD107880 36 human |
| 23 | 482 | 30.7 | 1499 | 6 | AR352504 | AR352504 Sequence |
| 24 | 482 | 30.7 | 1499 | 6 | BD005362 | BD005362 Protease |
| 25 | 482 | 30.7 | 1516 | 6 | BD107865 | BD107865 36 human |
| 26 | 482 | 30.7 | 1527 | 6 | BC008036 | BC008036 Homo sapi |
| 27 | 482 | 30.7 | 1570 | 6 | AR252580 | AR252580 Sequence |
| 28 | 482 | 30.7 | 1570 | 6 | AR528679 | AR528679 Sequence |
| 29 | 482 | 30.7 | 1570 | 6 | AX080829 | AX080829 Sequence |
| 30 | 482 | 30.7 | 1570 | 6 | AX403421 | AX403421 Sequence |
| 31 | 482 | 30.7 | 1570 | 6 | AX464322 | AX464322 Sequence |
| 32 | 482 | 30.7 | 1570 | 6 | AY359010 | AY359010 Homo sapi |
| 33 | 476 | 30.3 | 1476 | 6 | AR078184 | AR078184 Sequence |
| 34 | 476 | 30.3 | 1476 | 6 | AR137506 | AR137506 Sequence |
| 35 | 476 | 30.3 | 1476 | 6 | AR242358 | AR242358 Sequence |
| 36 | 476 | 30.3 | 1476 | 6 | BD082659 | BD082659 Keratinoc |
| 37 | 460 | 29.3 | 1504 | 6 | AR263823 | AR263823 Sequence |
| 38 | 449 | 28.6 | 673 | 6 | AX067353 | AX067353 Sequence |
| 39 | 315.4 | 20.1 | 498 | 11 | BV209338 | BV209338 KLU5.3564 |
| 40 | 236.6 | 15.1 | 321519 | 2 | AL714004 | AL714004 Homo sapi |
| 41 | 236 | 15.0 | 173559 | 9 | AC112458 | AC112458 Homo sapi |
| 42 | 234.4 | 14.9 | 174380 | 2 | AC016075 | AC016075 Homo sapi |
| 43 | 233.2 | 14.8 | 151880 | 9 | AP005716 | AP005716 Homo sapi |
| 44 | 231.2 | 14.7 | 79516 | 9 | AC004834 | AC004834 Homo sapi |
| 45 | 231.2 | 14.7 | 156949 | 9 | AL161652 | AL161652 Human DNA |

ALIGNMENTS

| | | | | | | |
|-----------------------|--------------------------------------|--|-----------|---------------|--------|-----------------|
| RESULT 1 | CQ788219 | CQ788219 | 11570 bp | DNA | linear | PAT 24-MAR-2004 |
| LOCUS | Sequence 2 from Patent WO2004021008. | | | | | |
| DEFINITION | CQ788219 | | | | | |
| ACCESSION | CQ788219.1 | GI:45723068 | | | | |
| VERSION | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | Homo sapiens (human) | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| REFERENCE | 1 | Diamonds, E. P. | | | | |
| AUTHORS | | Methods for detecting breast and ovarian cancer | | | | |
| TITLE | | Patent: WO 2004021008-A 2 11-MAR-2004; | | | | |
| JOURNAL | | MOUNT SINAI HOSPITAL (CA); Yousef, George (CA) | | | | |
| FEATURES | | Location/Qualifiers | | | | |
| source | | 1..11570 | | | | |
| ORIGIN | | /organism="Homo sapiens" | | | | |
| | | /mol_type="unassigned DNA" | | | | |
| | | /db_xref="taxon:9606" | | | | |
| Query Match | 100.0% | Score 1571; | DB 6; | Length 11570; | | |
| Best Local Similarity | 100.0% | Pred. No. 0; | | | | |
| Matches 1571; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | | |
| QY | 1 | AGGAGGATGGGGAATTGAGACAGAAACAAATTAAGTCCAGCGATGATTCTTA | 60 | | | |
| DB | 10000 | AGGAGGATGGGGAATTGAGACAGAAACAAATTAAGTCCAGCGATGATTCTTA | 10059 | | | |
| QY | 61 | TTGGAGTGTTCCTGCTTGAAGACATGGCAATCCAGAGACATTTTGGTTGTCA | 120 | | | |
| DB | 10060 | TTGGAGTGTTCCTGCTTGAAGACATGGCAATCCAGAGACATTTTGGTTGTCA | 10119 | | | |
| QY | 121 | CAACTATATGAGGGGCACTTATGCAATATGATGATGCAAGTGGCTTCAAC | 180 | | | |
| DB | 10120 | CAACTATATGAGGGGCACTTATGCAATATGATGATGCAAGTGGCTTCAAC | 10179 | | | |
| QY | 181 | ATGCTATGATGACACAGGAGGCTTCCACAAACATTAATCCAGATGCCAC | 240 | | | |
| DB | 10180 | ATGCTATGATGACACAGGAGGCTTCCACAAACATTAATCCAGATGCCAC | 10239 | | | |

QY 241 AGTCCAGATCGAGAACCTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGAGGAG 300
Db 10240 AGTGCCAGATCGAGAACCTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGAGGAG 10299
QY 301 GATTAAGATGGGTGGTGGAGAAATGGGAGAGAGGTGTGTCTCAATAGAGAAATTA 360
Db 10300 GATTAAGATGGGTGGTGGAGAAATGGGAGAGAGGTGTGTCTCAATAGAGAAATTA 10359
QY 361 GGCCTGACAGAGCTGAGAGGGAGAGTGAAGAGAGAGAGGGAGAGATACAGATTA 420
Db 10360 GGCCTGACAGAGCTGAGAGGGAGAGTGAAGAGAGAGAGGGAGAGATACAGATTA 10419
QY 421 GGGAGA CAGGCTGGAACAGAAATAGAGACGAAGATTTCCAGATGTGAGAGAGAGGCTCA 480
Db 10420 GGGAGA CAGGCTGGAACAGAAATAGAGACGAAGATTTCCAGATGTGAGAGAGAGGCTCA 10479
QY 481 CAGACCCCCCGGAATGATGTGTGACAA CAGGAATCTGAGAGAGAGAGATGAGTGTAG 540
Db 10480 CAGACCCCCCGGAATGATGTGTGACAA CAGGAATCTGAGAGAGAGAGATGAGTGTAG 10539
QY 541 AGTGA CAATGGGGTCTAAAGGTGAACCTTGGAGG CAGGACATGGTGCACCGCTGTA 600
Db 10540 AGTGA CAATGGGGTCTAAAGGTGAACCTTGGAGG CAGGACATGGTGCACCGCTGTA 10599
QY 601 ATCCCAACA CTTTGGAGGCTGAGAGGTGGGGAATCACTTGGAGG CAGAGATTGAGAC CAG 660
Db 10600 ATCCCAACA CTTTGGAGGCTGAGAGGTGGGGAATCACTTGGAGG CAGAGATTGAGAC CAG 10659
QY 661 CCTGGCAACATGTGTGAACCCCGTCTCTACAAAAAATA CA AAAAATTTAGCCGGGTGT 720
Db 10660 CCTGGCAACATGTGTGAACCCCGTCTCTACAAAAAATA CA AAAAATTTAGCCGGGTGT 10719
QY 721 GGTATGAGACACCTGTATGTCACAGTACTTGGAGGCTGAGAGAGAGATGCTTGAC 780
Db 10720 GGTATGAGACACCTGTATGTCACAGTACTTGGAGGCTGAGAGAGATGCTTGAC 10779
QY 781 CCGGAGATGAGAGCTGAGTGAAGTGAAGTCAAGGCACATGCGCTCAACCTGGGACACA 840
Db 10780 CCGGAGATGAGAGCTGAGTGAAGTGAAGTCAAGGCACATGCGCTCAACCTGGGACACA 10839
QY 841 GAGTAAAGATCTCCATCTCAAAAAAATAAGCTGAGATTGAGAGTGAATTTAATAACATT 900
Db 10840 GAGTAAAGATCTCCATCTCAAAAAAATAAGCTGAGATTGAGAGTGAATTTAATAACATT 10899
QY 901 CTCCTCTCTCTCTTGTGCTGTGTCTGCATCTCTGCTTTTTCGACATTTCTTATCT 960
Db 10900 CTCCTCTCTCTCTTGTGCTGTGTCTGCATCTCTGCTTTTTCGACATTTCTTATCT 10959
QY 961 CTGTACTTTCATCTCTGTGTCTGTGTCTGCATCTGCTTCTGCATCTATGAGGATCTTG 1020
Db 10960 CTGTACTTTCATCTCTGTGTCTGTGTCTGCATCTGCTTCTGCATCTATGAGGATCTTG 11019
QY 1021 GGTCTCTCATGTCTCTCTTCTGCCCATTTTTCGACATCTCTGCTCTCTCAATGCCCCCTT 1080
Db 11020 GGTCTCTCATGTCTCTCTTCTGCCCATTTTTCGACATCTCTGCTCTCTCAATGCCCCCTT 11079
QY 1081 TCTCTCTGAGAGGATTTCTGGGGGGGCTGTGTGTGCAATGGGCTCCCTGACAGGAGATC 1140
Db 11080 TCTCTCTGAGAGGATTTCTGGGGGGGCTGTGTGTGCAATGGGCTCCCTGACAGGAGATC 11139
QY 1141 GTGTCTCTGGGAGATTAACCTTGTGCCCCGCCCAACAGACCGGGGTCTTACAGAACTTC 1200
Db 11140 GTGTCTCTGGGAGATTAACCTTGTGCCCCGCCCAACAGACCGGGGTCTTACAGAACTTC 11199
QY 1201 TGGCAATTCACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGGA 1260
Db 11200 TGGCAATTCACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGGA 11259
QY 1261 CTCAGACACCGGATCCCACTGCTGAGAGGAGAGCCCTGACATCTCTTTCAGAGCCCT 1320
Db 11260 CTCAGACACCGGATCCCACTGCTGAGAGGAGAGCCCTGACATCTCTTTCAGAGCCCT 11319

QY 1321 CATTCCTCCAGAGATGTGGAATGTTTCATCTCTCAGAGCCCTGACCCCATGTCCT 1380
Db 11320 CATTCCTCCAGAGATGTGGAATGTTTCATCTCTCAGAGCCCTGACCCCATGTCCT 11379
QY 1381 GGAATCAGGGTGTGCTTCCCCCAATTTGGGCTGACCGGTGTCTCTAGTTGAACCTTGG 1440
Db 11380 GGAATCAGGGTGTGCTTCCCCCAATTTGGGCTGACCGGTGTCTCTAGTTGAACCTTGG 11439
QY 1441 AACCAATTTCAAAACGTCTCAGAGGCGGGGTTGCGTCTCAATCTCCCTGGGACCTTCA 1500
Db 11440 AACCAATTTCAAAACGTCTCAGAGGCGGGGTTGCGTCTCAATCTCCCTGGGACCTTCA 11499
QY 1501 TCCCTAGCTCAGAGGCCCATCCCTTCTCTGACGCTTGAACCCAAATTTAGTCCAGAAAT 1560
Db 11500 TCCCTAGCTCAGAGGCCCATCCCTTCTCTGACGCTTGAACCCAAATTTAGTCCAGAAAT 11559
QY 1561 AAATGAGAG 1571
Db 11560 AAATGAGAG 11570

RESULT 2
CQ874885
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent WO2004077060.
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 Diamandis, E.P. and Petraki, C.
TITLE Assay for detection of renal cell carcinoma
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;
Mount Sinai Hospital (CA)
FEATURES
source
1..11570
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1571; DB 6; Length 11570;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGTAATGGGAATTGAAGACAGAAACAAATTAATGTCACAGGATGATTTCTA 60
Db 10000 AGGAGGTAATGGGAATTGAAGACAGAAACAAATTAATGTCACAGGATGATTTCTA 10059
QY 61 TTGGAGATGATTTCTGCCCCCTAGAAACACTGGCATATCCAGAGACATTTTGTGTGCA 120
Db 10060 TTGGAGATGATTTCTGCCCCCTAGAAACACTGGCATATCCAGAGACATTTTGTGTGCA 10119
QY 121 CAATATATGAGAGGGGCAATTTACTGCAACTAATGATGATGCAAGTGTGCTTCAAC 180
Db 10120 CAATATATGAGAGGGGCAATTTACTGCAACTAATGATGATGCAAGTGTGCTTCAAC 10179
QY 181 ATGCTATATGACACCGGACGCTTCCAAACAAACATTAATCCAGCTTCAGATGCCAC 240
Db 10180 ATGCTATATGACACCGGACGCTTCCAAACAAACATTAATCCAGCTTCAGATGCCAC 10239
QY 241 AGTGCCAGATGAGAAACCTTCATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 300
Db 10240 AGTGCCAGATGAGAAACCTTCATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 10299
QY 301 GATTAAGATGGGTGGTGGAGAAATGGGAGAGAGGTGTGTGTCAGTAAAGAAATTA 360
Db 10300 GATTAAGATGGGTGGTGGAGAAATGGGAGAGAGGTGTGTGTCAGTAAAGAAATTA 10359
QY 361 GGCCTGACAGGCTGAGAGGGAGAGTGAAGAGAGAGAGAGGCGGAGAGATACAGATGA 420

Db 10360 GGGCTGCACAGGCTGGAGGGGAGGTAGAGAGAAAGGAGGGGAGAGATTACACATGA 10419
Qy 421 GGGAGACAGGCTGGAAACAGAAAGTACAGAAAGATTGAGATGTGAGAGAAAGGTCA 480
Db 10420 GGGAGACAGGCTGGAAACAGAAAGTACAGAAAGATTGAGATGTGAGAGAAAGGTCA 10479
Qy 481 CAGACCCCGGAGAAAGTGTGAGCAACAGAAATCTGGAAGAGAAAGTGTGAGAG 540
Db 10480 CAGACCCCGGAGAAAGTGTGAGCAACAGAAATCTGGAAGAGAAAGTGTGAGAG 10539
Qy 541 AGTGAACAAATGGGGTCTAAAGTTGAATTGAGAGGCGAGCATGTGGCTCAGCCCTGA 600
Db 10540 AGTGAACAAATGGGGTCTAAAGTTGAATTGAGAGGCGAGCATGTGGCTCAGCCCTGA 10599
Qy 601 ATCCCAACACTTTGAGAGGCTGAGAGTGGGCAATCACTTGAAGCCAGAGATTGAGACAG 660
Db 10600 ATCCCAACACTTTGAGAGGCTGAGAGTGGGCAATCACTTGAAGCCAGAGATTGAGACAG 10659
Qy 661 CCTGGCAACATGTTGAAAACCCGCTCTCAAAAAAATCAAAAAATTAGCCGGGTCT 720
Db 10660 CCTGGCAACATGTTGAAAACCCGCTCTCAAAAAAATCAAAAAATTAGCCGGGTCT 10719
Qy 721 GGTGATGACACCTGTAGTACAGCTACTTGGAGGCTGAGAGAGAGAAATGCTTGAAC 780
Db 10720 GGTGATGACACCTGTAGTACAGCTACTTGGAGGCTGAGAGAGAAATGCTTGAAC 10779
Qy 781 CCGGAGATGAGAGGCTGAGAGTGAAGGTCAAGGCCACTGGCTCCAACTGGGCAACA 840
Db 10780 CCGGAGATGAGAGGCTGAGAGTGAAGGTCAAGGCCACTGGCTCCAACTGGGCAACA 10839
Qy 841 GAGTAAGACTCCCATCTCAAAAAAAGCTGATTTGAGAGTAAATATTAATCAAT 900
Db 10840 GAGTAAGACTCCCATCTCAAAAAAAGCTGATTTGAGAGTAAATATTAATCAAT 10899
Qy 901 CTCCCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGATTTCTTCAATCT 960
Db 10900 CTCCCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGATTTCTTCAATCT 10959
Qy 961 CTGTACTTTCATCTCTGTGTCTGTGCTCCATCTGTCTTCAATCTTATGAGCAATCTCTG 1020
Db 10960 CTGTACTTTCATCTCTGTGTCTGTGCTCCATCTGTCTTCAATCTTATGAGCAATCTCTG 11019
Qy 1021 GGTCTCTCAATGTCTCTTCTGCCCACCTTGGCCACATCTGCTCTCTCAATGCCCCCTT 1080
Db 11020 GGTCTCTCAATGTCTCTTCTGCCCACCTTGGCCACATCTGCTCTCTCAATGCCCCCTT 11079
Qy 1081 TCTCTCTGAGGGGTGATTTCTGGGGGGGCTGTGGTCTGCAATGGCTCCCTGACAGGACTC 1140
Db 11080 TCTCTCTGAGGGGTGATTTCTGGGGGGGCTGTGGTCTGCAATGGCTCCCTGACAGGACTC 11139
Qy 1141 GTGTCTCTGGGAGATTACCTTGTGTGCCGGCCAAACAGACCGGGGTGTCTACAGAACTTC 1200
Db 11140 GTGTCTCTGGGAGATTACCTTGTGTGCCGGCCAAACAGACCGGGGTGTCTACAGAACTTC 11199
Qy 1201 TGTAAATTACCAAGTGTGATTCAGAAACATTCAGGGCCAACTCTGAGTCAATCCCAAGA 1260
Db 11200 TGTAAATTACCAAGTGTGATTCAGAAACATTCAGGGCCAACTCTGAGTCAATCCCAAGA 11259
Qy 1261 CTGAGACACCGGATCCCCACCTGCTGAGAGGACAGCCCTGACACTCTTTGAGACCTT 1320
Db 11260 CTGAGACACCGGATCCCCACCTGCTGAGAGGACAGCCCTGACACTCTTTGAGACCTT 11319
Qy 1321 CATTCCTTCCAGAGATGTGAGAAATGTTCACTCTTCAGACCCCTGAACCCCATGTCTCT 1380
Db 11320 CATTCCTTCCAGAGATGTGAGAAATGTTCACTCTTCAGACCCCTGAACCCCATGTCTCT 11379
Qy 1381 GGAAGTCAAGGTTCTCTTCCCCCAATTGGGCTGACCTGTCTCTAGTTGAAACCTCTGG 1440
Db 11380 GGAAGTCAAGGTTCTCTTCCCCCAATTGGGCTGACCTGTCTCTAGTTGAAACCTCTGG 11439
Qy 1441 AACAAATTTCCAAATGTCAGAGGGGGGGGTGCGTCTCAATCTCCCTGGGGGCACTTCA 1500

Db 11440 AACAAATTTCCAAATGTCAGAGGGGGGGTGTGCTCAATCTCCCTGGGGGCACTTCA 11499
Qy 1501 TCTCAAGCTCAGAGGCCCATTCCTTCTCTGAGCTGTGACCCCAATTTAGTCCCAAAAT 1560
Db 11500 TCTCAAGCTCAGAGGCCCATTCCTTCTCTGAGCTGTGACCCCAATTTAGTCCCAAAAT 11559
Qy 1561 AAAGTGAAG 1571
Db 11560 AAAGTGAAG 11570

RESULT 3
COB74961 11570 bp DNA linear PAT 27-SEP-2004
LOCUS COB74961
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION COB74961 GI:52748060
VERSION COB74961.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Diamandis, E.P.
AUTHORS Multiple marker assay for detection of ovarian cancer
TITLE Patent: WO 2004075713-A 5 10-SEP-2004;
JOURNAL MOUNT SINAI HOSPITAL CORPORATION (CA)
MOUNT SINAI HOSPITAL
FEATURES
source
1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1571; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGAGGTATGGGAAATTTGAAGACAGAAACACAAATTTAGTCCAAACGAATGATTTCTTA 60
Db 10000 AGGAGGTATGGGAAATTTGAAGACAGAAACACAAATTTAGTCCAAACGAATGATTTCTTA 10059
Qy 61 TTGGAGTGAATTTCTGCCCCCTAGAAACATGCGCAATACAGAGAGACATTTTGGTTGTCA 120
Db 10060 TTGGAGTGAATTTCTGCCCCCTAGAAACATGCGCAATACAGAGAGACATTTTGGTTGTCA 10119
Qy 121 CAATATATGAGGGGCACTTACTGTGCACTAAATGATAGTCAAGTGTCTTCAAC 180
Db 10120 CAATATATGAGGGGCACTTACTGTGCACTAAATGATAGTCAAGTGTCTTCAAC 10179
Qy 181 ATGCTATGATGACACGGGAGGCTTCAACAAACATTTATTCAGCTTCAAGTCCAC 240
Db 10180 ATGCTATGATGACACGGGAGGCTTCAACAAACATTTATTCAGCTTCAAGTCCAC 10239
Qy 241 AGTCCCAATGAGGAAACCTCATCAGGGGTGGAACCGATTTTTCAGAAAGGAG 300
Db 10240 AGTCCCAATGAGGAAACCTCATCAGGGGTGGAACCGATTTTTCAGAAAGGAG 10299
Qy 301 GTATTAAGATGGGTTGGTGAAGATGGGAAAGAGTGTGTCCAGTAAAGAAATTA 360
Db 10300 GTATTAAGATGGGTTGGTGAAGATGGGAAAGAGTGTGTCCAGTAAAGAAATTA 10359
Qy 361 GGCCTGCAAGGCTGAGAGGGGAGAGTGAAGAGAAAGGAGGCGGAGATACATGA 420
Db 10360 GGCCTGCAAGGCTGAGAGGGGAGAGTGAAGAGAAAGGAGGCGGAGATACATGA 10419
Qy 421 GGGAGACAGGCTGGAACAGAAATGAGAGCGAAGATTGAGATGTGAGAGAGAGGCTCA 480
Db 10420 GGGAGACAGGCTGGAACAGAAATGAGAGCGAAGATTGAGATGTGAGAGAGAGGCTCA 10479
Qy 481 CAGACCCCGGAGAAATGATGTGACAAACAGAAATCTGGAAGAGAGAGATGAGTGAAG 540
Db 10480 CAGACCCCGGAGAAATGATGTGACAAACAGAAATCTGGAAGAGAGAGATGAGTGAAG 10539

| | | | |
|----|-------|---|--------|
| QY | 541 | AGTGAACAAATGGGGGTCTAAAGATTGAACTTGGAGGCGCAGGCAATGGTGGCTCAACGGCTGTA | 600 |
| Db | 10540 | AGTGAACAAATGGGGTCTAAAGATTGAACTTGGAGGCGCAGGCAATGGTGGCTCAACGGCTGTA | 105999 |
| QY | 601 | ATCCCAACACTTTTGGAGGCTGAAGGTGGCGAATCACTTGAAGCCAGAGATTGCGACGAG | 660 |
| Db | 10600 | ATCCCAACACTTTTGGAGGCTGAAGGTGGCGAATCACTTGAAGCCAGAGATTGCGACGAG | 106599 |
| QY | 661 | CTGGGCCAACAATGGTGAACCCCGTCTCTACAAAAAAATACAAAAAATTAAGCCGGGTCT | 720 |
| Db | 10660 | CTGGGCCAACAATGGTGAACCCCGTCTCTACAAAAAATTAAGCCGGGTCT | 107199 |
| QY | 721 | GGTATGGAACACTGTATGACACAGCTACTTGGAGGCTGAGGCAAGAGATTGCTTGAAC | 780 |
| Db | 10720 | GGTATGGAACAACCTGTATGACACAGCTACTTGGAGGCTGAGGCAAGAGATTGCTTGAAC | 107799 |
| QY | 781 | CCGGGAATAGAGGCTTCATATGAGTGAAGGTCAAGGCCATCGCGCTCCAACTGGGCAACA | 840 |
| Db | 10780 | CCGGGAATAGAGGCTTCATATGAGTGAAGGTCAAGGCCATCGCGCTCCAACTGGGCAACA | 108399 |
| QY | 841 | GAGTAAGAATCTCAATCTCAAAAAAATAAAGCTGATTAATGGAGTAATTAATTAACATT | 900 |
| Db | 10840 | GAGTAAGAATCTCAATCTCAAAAAAATAAAGCTGATTAATGGAGTAATTAATTAACATT | 108999 |
| QY | 901 | CTCCCTCTCTCTCTCTTTGGCCTGTGTCTCCATCTGTGTCTTTTTCGACATTTCTTCATCT | 960 |
| Db | 10900 | CTCCCTCTCTCTCTCTTTGGCCTGTGTCTCCATCTGTGTCTTTTTCGACATTTCTTCATCT | 109599 |
| QY | 961 | CTGTACTTTCCATCTCTGTGTGTCTGTGTCCCATCTGTCTTCCATCTAATAGGGCAATCTCG | 1020 |
| Db | 10960 | CTGTACTTTCCATCTCTGTGTGTCTGTGTCCCATCTGTCTTCCATCTAATAGGGCAATCTCG | 110199 |
| QY | 1021 | GGTCTCTCATGTCTCTTTCGCCCATTTTGGCCACATCTGTGCTCTCTCATGCCCCCTT | 1080 |
| Db | 11020 | GGTCTCTCATGTCTCTTTCGCCCATTTTGGCCACATCTGTGCTCTCTCATGCCCCCTT | 110799 |
| QY | 1081 | TCTCTCTTGAAGGGTATTTCTGGGGGGGCTGTGTGTCTGCATATGGCTTCCGAGGGACCTC | 1140 |
| Db | 11080 | TCTCTCTTGAAGGGTATTTCTGGGGGGGCTGTGTGTCTGCATATGGCTTCCGAGGGACCTC | 111399 |
| QY | 1141 | GTTGTCCTGGGGAAGTTATCCCTTGTGCCCCGCCCAACAGACCGGGGTGTACACGAACCTC | 1200 |
| Db | 11140 | GTTGTCCTGGGGAAGTTATCCCTTGTGCCCCGCCCAACAGACCGGGGTGTACACGAACCTC | 111399 |
| QY | 1201 | TGCAAGTTCAACAAATGATTCAGAGAAACCATCAGGCGCAACTCTGAATCATCCAGGA | 1260 |
| Db | 11200 | TGCAAGTTCAACAAATGATTCAGAGAAACCATCAGGCGCAACTCTGAATCATCCAGGA | 112599 |
| QY | 1261 | CTGAGCAACACCGGATCCCAACCTGTCGACAGGGGACAGCCCTGACATCTCTTTCAGACCTT | 1320 |
| Db | 11260 | CTGAGCAACACCGGATCCCAACCTGTCGACAGGGGACAGCCCTGACATCTCTTTCAGACCTT | 113199 |
| QY | 1321 | CATTCTTCCAGAGATGTGAAGATGTTCATCTCTGACAGCCCTGACCCCAATGTCTCCT | 1380 |
| Db | 11320 | CATTCTTCCAGAGATGTGAAGATGTTCATCTCTGACAGCCCTGACCCCAATGTCTCCT | 113799 |
| QY | 1381 | GGAATCAGGGTCTGTCTTCCCCCAATTTGGGCTACCGGTGTCTCTTAAGTTGAACCTGTGG | 1440 |
| Db | 11380 | GGAATCAGGGTCTGTCTTCCCCCAATTTGGGCTACCGGTGTCTCTTAAGTTGAACCTGTGG | 114399 |
| QY | 1441 | AACAAATTTCAAAAATGTCAAGGGCGGGGGTTCGATCTCAATCTCCCTGAGGCACTTTCA | 1500 |
| Db | 11440 | AACAAATTTCAAAAATGTCAAGGGCGGGGGTTCGATCTCAATCTCCCTGAGGCACTTTCA | 114999 |
| QY | 1501 | TCTCTCAAGCTCAGGGCCCATCTCCTTCTCTGACAGCTTGACCCAAATTTAAGTCCAGAAAT | 1560 |
| Db | 11500 | TCTCTCAAGCTCAGGGCCCATCTCCTTCTCTGACAGCTTGACCCAAATTTAAGTCCAGAAAT | 115599 |
| QY | 1561 | AAACTGAGAG 1571 | |
| Db | 11560 | AAACTGAGAG 11570 | |

| LOCUS | AF135028 | 11570 bp | DNA | linear | PRI 27-JUN-2000 |
|-----------------------|---|------------|-----|--------|-----------------|
| DEFINITION | Homo sapiens kallikrein-like protein 2 KUK-L2 gene, complete cds. | | | | |
| ACCESSION | AF135028 | | | | |
| VERSION | AF135028.1 | GI:4589282 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Homo sapiens (human) | | | | |
| REFERENCE | Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| TITLE | 1 (bases 1 to 11570) | | | | |
| AUTHORS | Yousef, G.M. and Diamandis, E.P. | | | | |
| TITLE | The new kallikrein-like gene, KUK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation | | | | |
| JOURNAL | J. Biol. Chem. 274 (53), 37511-37516 (1999) | | | | |
| MEDLINE | 20076408 | | | | |
| PUBMED | 10608802 | | | | |
| REFERENCE | 2 (bases 1 to 11570) | | | | |
| AUTHORS | Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obiezu, C.V. | | | | |
| TITLE | The new human kallikrein gene family: implications in | | | | |
| JOURNAL | cachinogenesis | | | | |
| MEDLINE | 1121728 | | | | |
| PUBMED | 10675891 | | | | |
| REFERENCE | 3 (bases 1 to 11570) | | | | |
| AUTHORS | Yousef, G.M., Luo, L.Y. and Diamandis, E.P. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada | | | | |
| FEATURES | | | | | |
| source | Location/Qualifiers | | | | |
| | 1..11570 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| | /chromosome="19" | | | | |
| | /map="19q13.3-q13.4" | | | | |
| | join(2101..2131,2210..2293,4762..5023,5763..6019,6105..6238,11092..11570) | | | | |
| | /product="kallikrein-like protein 2 KUK-L2" | | | | |
| | join(2221..2293,4762..5023,5763..6019,6105..6238,11092..11247) | | | | |
| | /codon_start=1 | | | | |
| | /product="kallikrein-like protein 2 KUK-L2" | | | | |
| | /protein_id="AA026429.1" | | | | |
| | /db_xref="GI:4589283" | | | | |
| | /translation="MATARPMMWVICALITALLIGTEVHLANDVSCDPSNTVPSGNSQDVGAGGEBARSDSSSRILVINGSDCMHTOPQALLRPQLXGVALVHPQW LLSAAARKKVFRVRLGHSLSPVYSGQMGQGVKSLIPPEYSHRGHNDMLIKLN RRIAPRPDVRPIVSHSPGASAGKCLVSGMGTTKSPQVHPVYLQCLNLSVLSORCE DAPROLDIPMPCAGDKAGRDSQQDSSGPVVCNLSIQGLVSWGDIPCARPRRPGVYT NCKFTKRWIOETIQANS" | | | | |
| ORIGIN | | | | | |
| Query Match | 100.0%; Score 1571; DB 9; Length 11570; | | | | |
| Beat Local Similarity | 100.0%; Pred. No. 0; | | | | |
| Matches 1571; | Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY | 1 AGGAGGATATGGGGAATTGAAGACAGAGAAACAAATTAATGTCACAGGATGATTTCTA 60 | | | | |
| Db | 10000 AGGAGGATATGGGGAATTGAAGACAGAGAAACAAATTAATGTCACAGGATGATTTCTA 10059 | | | | |
| QY | 61 TTGGAGATGATTCCTGCCCTCAGAAACACCTGCGAATACAGAGACATTTTGGTGTCA 120 | | | | |
| Db | 10060 TTGGAGATGATTCCTGCCCTCAGAAACACCTGCGAATACAGAGACATTTTGGTGTCA 10119 | | | | |
| QY | 121 CAACTATATGAGAGGCATTTACGCAACTAATGATAGATCCCAAGTGTGCTTCAAC 180 | | | | |
| Db | 10120 CAACTATATGAGAGGCATTTACGCAACTAATGATAGATCCCAAGTGTGCTTCAAC 10179 | | | | |
| QY | 181 ATGCTATGATGACACGGCAGCCTTCACACAAACCATTTATTCAGCTTCAGATGCCAC 240 | | | | |

```

Db      10180 ATGTATGATGACACCGGAGGCTCCACAACAAATATTCAGCTTGAGATGCCAC 10239
Qy      241 AGTCCCAAGATCGAGGAACCTCATCCAGGGGCTGAGAACCGTATTTTGGCAGAGGAG 300
Db      10240 AGTCCCAAGATCGAGGAACCTCATCCAGGGGCTGAGAACCGTATTTTGGCAGAGGAG 10299
Qy      301 GTATTAAGATGGGTGGTGGAGAAATGGGGAAGGAAGGTGTGTCTCAATAGAGAAATTA 360
Db      10300 GTATTAAGATGGGTGGTGGAGAAATGGGGAAGGAAGGTGTGTCTCAATAGAGAAATTA 10359
Qy      361 GGCCTGACAGGCTGAGAGGGGAGAGTGAAGAGAAAGGAGCCGAGAGATACAGATGA 420
Db      10360 GGCCTGACAGGCTGAGAGGGGAGAGTGAAGAGAAAGGAGCCGAGAGATACAGATGA 10419
Qy      421 GGGAGACAGGCTGGAACAGAAAGTGAAGACAGAGATTCAGATGTGAGAGAGGATCA 480
Db      10420 GGGAGACAGGCTGGAACAGAAAGTGAAGACAGAGATTCAGATGTGAGAGAGGATCA 10479
Qy      481 CAGACCCCGCGAAATGATGTGTGAGCAACGGAATCGGAAGAGAGAGATGAGTGGAG 540
Db      10480 CAGACCCCGCGAAATGATGTGTGAGCAACGGAATCGGAAGAGAGAGATGAGTGGAG 10539
Qy      541 AGTGAACAAATGGGCTTAAAGTTGAACTTGAAGCCAGGCGATGAGCTCACGCTGTA 600
Db      10540 AGTGAACAAATGGGCTTAAAGTTGAACTTGAAGCCAGGCGATGAGCTCACGCTGTA 10599
Qy      601 ATCCCAACACTTTGAGAGGCTGAGAGGTGGCGCAATCACTTGAAGCCAGAGATTGAGAC 660
Db      10600 ATCCCAACACTTTGAGAGGCTGAGAGGTGGCGCAATCACTTGAAGCCAGAGATTGAGAC 10659
Qy      661 CTGGGCAACATGGTGAACCCGCTCTCAAAAAAATTAACAAATTAACCGGCTGT 720
Db      10660 CTGGGCAACATGGTGAACCCGCTCTCAAAAAAATTAACAAATTAACCGGCTGT 10719
Qy      721 GGTGATGACACCTGTAGTACAGCTTCTTGAAGGCTGAGGAGAGAAATGGCTTGAAC 780
Db      10720 GGTGATGACACCTGTAGTACAGCTTCTTGAAGGCTGAGGAGAGAAATGGCTTGAAC 10779
Qy      781 CCGGAGATGAGGCTGACAGTGAAGTCAAGGCACTGCGCTCCAACTGGGCAACA 840
Db      10780 CCGGAGATGAGGCTGACAGTGAAGTCAAGGCACTGCGCTCCAACTGGGCAACA 10839
Qy      841 GAGTAACACTCCATCTCAAAAAAAGCTGGAATTTGAAGTAATTAATAAAT 900
Db      10840 GAGTAACACTCCATCTCAAAAAAAGCTGGAATTTGAAGTAATTAATAAAT 10899
Qy      901 CTCCTCTCTCTCTTGGCTGTGCTCCATCTCTGCTTTTCTGCAATTTCTTCATCT 960
Db      10900 CTCCTCTCTCTCTTGGCTGTGCTCCATCTCTGCTTTTCTGCAATTTCTTCATCT 10959
Qy      961 CTGACTTTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db      10960 CTGACTTTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11019
Qy      1021 GGTCTCTCAATGCTCTTCCCTTCCCACTTGGCAATCTGCGCTCTTCAATGCCCTT 1080
Db      11020 GGTCTCTCAATGCTCTTCCCTTCCCACTTGGCAATCTGCGCTCTTCAATGCCCTT 11079
Qy      1081 TCTCTCTGACGGGTGATTTCTGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db      11080 TCTCTCTGACGGGTGATTTCTGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11139
Qy      1141 GTGTCTGGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db      11140 GTGTCTGGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11199
Qy      1201 TGCAGATTCAACAGTGAATTCAGAGAAACATTCAGGCAACTCTGAGTCAATCCAGGA 1260
Db      11200 TGCAGATTCAACAGTGAATTCAGAGAAACATTCAGGCAACTCTGAGTCAATCCAGGA 11259
Qy      1261 CTCAGCACCGGATCCCACTGCTGACAGGAGACCCCTGACACTCTTTGAGACCTT 1320

```

```

Db      11260 CTCAGCACCGGATCCCACTGCTGACAGGAGACCCCTGACACTCTTTGAGACCT 11319
Qy      1321 CATTCCTCCAGAGATGTGAAATGTTATCTCCAGGCCCTGACCCCATGTCCT 1380
Db      11320 CATTCCTCCAGAGATGTGAAATGTTATCTCCAGGCCCTGACCCCATGTCCT 11379
Qy      1381 GGACTCAGGCTGTCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTTGGG 1440
Db      11380 GGACTCAGGCTGTCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTTGGG 11439
Qy      1441 AACAAATTCCAAACTGTCCAGGCGCGGGGTTCGCTCAATCTCCCTGGGCACTTCA 1500
Db      11440 AACAAATTCCAAACTGTCCAGGCGCGGGGTTCGCTCAATCTCCCTGGGCACTTCA 11499
Qy      1501 TCCTCAAGCTCAGGGCCCAATCCCTCTCTGACGCTGACCCAAATTAAGTCCAGAAAT 1560
Db      11500 TCCTCAAGCTCAGGGCCCAATCCCTCTCTGACGCTGACCCAAATTAAGTCCAGAAAT 11559
Qy      1561 AAACGTGAGAG 1571
Db      11560 AAACGTGAGAG 11570

RESULT 5
AC011483/c 107487 bp DNA linear PRI 01-JUL-2002
LOCUS AC011483
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
ACCESSION AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Unpublished
2 (bases 1 to 107487)
DOE Joint Genome Institute.
REFERENCE Direct Submission
Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE Direct Submission
Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jul 1, 2002 this sequence version replaced gi:14971176.
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1. 107487
/morganism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C22"
64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."

ORIGIN
Query Match 100.0%; Score 1571; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 1 AGGAGGTATGGGAAATTGAAAGACAGAAACAAATTAATGCTCAAGCATGATTTCTA 60
DB 50220 AGGAGGTATGGGAAATTGAAAGACAGAAACAAATTAATGCTCAAGCATGATTTCTA 50161
QY 61 TTGGAGTGTATTTGCCCTTGAAGACCTGGCAATACAGAGACATTTTGGTTGTC 120
DB 50160 TTGGAGTGTATTTGCCCTTGAAGACCTGGCAATACAGAGACATTTTGGTTGTC 50101
QY 121 CAACATATGAGAGGGGCAATTACTGGCAACTAATGATAGATGACAGTGGCTTCAAC 180
DB 50100 CAACATATGAGAGGGGCAATTACTGGCAACTAATGATAGATGACAGTGGCTTCAAC 50041
QY 181 ATGCTATGATGACACGCGAGGCTCCACAAACCAATATCAGCTTCAGATGCCAC 240
DB 50040 ATGCTATGATGACACGCGAGGCTCCACAAACCAATATCAGCTTCAGATGCCAC 49981
QY 241 AGTGCCAGATGAGAGAACCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGAG 300
DB 49980 AGTGCCAGATGAGAGAACCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGAG 49921
QY 301 GTATAGATGAGGTGTGGAGATGGGGAGGAAGGTGTCTCCAGTAAAGAAATTA 360
DB 49920 GTATAGATGAGGTGTGGAGATGGGGAGGAAGGTGTCTCCAGTAAAGAAATTA 49861
QY 361 GGCCTGACAGGCTGAGGAGGAGATGAGAGAGAGAGGAGCGAGAGATACAGATGA 420
DB 49860 GGCCTGACAGGCTGAGGAGGAGATGAGAGAGAGAGGAGCGAGAGATACAGATGA 49801
QY 421 GGGAGACAGGCTGGAACAGAAATGAGACAGAAATTCAGATGTGAGAGAGGCTCA 480
DB 49800 GGGAGACAGGCTGGAACAGAAATGAGACAGAAATTCAGATGTGAGAGAGGCTCA 49741
QY 481 CAGACCCCCCGAAATGATGTGTGACACAGGAATCTGGAAGAGAGATGAGAGTGGAG 540
DB 49740 CAGACCCCCCGAAATGATGTGTGACACAGGAATCTGGAAGAGAGATGAGAGTGGAG 49681
QY 541 AGTGACAAATGGGGTCTAAAGTTGAACTTGGAGGCGAGGATGTGGCTCAAGCTCTGA 600
DB 49680 AGTGACAAATGGGGTCTAAAGTTGAACTTGGAGGCGAGGATGTGGCTCAAGCTCTGA 49621
QY 601 ATCCCAACACTTTGGAGGCTGAGAGTGGCGAATCACTTGAAGCGAGAGATTCAGACAG 660
DB 49620 ATCCCAACACTTTGGAGGCTGAGAGTGGCGAATCACTTGAAGCGAGAGATTCAGACAG 49561
QY 661 CTGGGCAACATGATGTAAACCCGCTCTACAAATAAATAAATAAATAAATAAATAAATAA 720
DB 49560 CTGGGCAACATGATGTAAACCCGCTCTACAAATAAATAAATAAATAAATAAATAAATAA 49501
QY 721 GGTGATGACACCTGTAGTCAACGTAATTTGGAGGCTGAGGAGAGAGATTTGCTTGAAC 780
DB 49500 GGTGATGACACCTGTAGTCAACGTAATTTGGAGGCTGAGGAGAGAGATTTGCTTGAAC 49441
QY 781 CCGGAGATGAGAGGCTGAGTGAAGCTGAGAGTCAAGGCCACCTGGCTCAACTGGGCAACA 840
DB 49440 CCGGAGATGAGAGGCTGAGTGAAGCTGAGAGTCAAGGCCACCTGGCTCAACTGGGCAACA 49381
QY 841 GAGTAAAGATCCCATCTCAAAAAAATAAAGCTGAGATTTGGAGGAATAATTAATTAACAT 900
DB 49380 GAGTAAAGATCCCATCTCAAAAAAATAAAGCTGAGATTTGGAGGAATAATTAATTAACAT 49321
QY 901 CTCCCT 960
DB 49320 CTCCCT 49261
QY 961 CTGTATCTTTCATCTCTGT 1020
DB 49260 CTGTATCTTTCATCTCTGT 49201
QY 1021 GGTCTCTCAATGTCT 1080
DB 49200 GGTCTCTCAATGTCT 49141

QY 1081 TCTCTCTGACAGGGTGAATTTCTGGGGGGGCTGTGGTCTGCAATGCTCCCTGACGGACTC 1140
DB 49140 TCTCTCTGACAGGGTGAATTTCTGGGGGGGCTGTGGTCTGCAATGCTCCCTGACGGACTC 49081
QY 1141 GTGTCTGGGGAGATTAACCTTGTGTGCGGCCCAACAGACCGGGGTGTCTACACGAACCTC 1200
DB 49080 GTGTCTGGGGAGATTAACCTTGTGTGCGGCCCAACAGACCGGGGTGTCTACACGAACCTC 49021
QY 1201 TGCAAGTTACCAAGTGAATTCAGAGAAACATCCAGGGCCCACTCTGAGTCAATCCAGGA 1260
DB 49020 TGCAAGTTACCAAGTGAATTCAGAGAAACATCCAGGGCCCACTCTGAGTCAATCCAGGA 48961
QY 1261 CTCAGACACCGGACATCCCACTCTGAGGAGACCCCTGACATCTCTTTCAGACCT 1320
DB 48960 CTCAGACACCGGACATCCCACTCTGAGGAGACCCCTGACATCTCTTTCAGACCT 48901
QY 1321 CATTCCTTCCAGAGATGTTGAAATGTTTCAATCTCTCCAGCCCTGACCCCAATGTCTCT 1380
DB 48900 CATTCCTTCCAGAGATGTTGAAATGTTTCAATCTCTCCAGCCCTGACCCCAATGTCTCT 48841
QY 1381 GGAAGTCAAGGGTGTGCTCCGCCCAATGGGGCTGACCGGTCTCTGATGTAACCTGGG 1440
DB 48840 GGAAGTCAAGGGTGTGCTCCGCCCAATGGGGCTGACCGGTCTCTGATGTAACCTGGG 48781
QY 1441 AACAAATTTCCAAAGTGTCCAGGGGCGGGGTTGCTTCAATCTCTGAGGCACTTTCA 1500
DB 48780 AACAAATTTCCAAAGTGTCCAGGGGCGGGGTTGCTTCAATCTCTGAGGCACTTTCA 48721
QY 1501 TCTCTCAAGTCAAGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTTATGTCAGGAAT 1560
DB 48720 TCTCTCAAGTCAAGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTTATGTCAGGAAT 48661
QY 1561 AAACCTGAGAG 1571
DB 48660 AAACCTGAGAG 48650

RESULT 6
AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
LOCUS AF243527
DEFINITION Homo sapiens serine protease gene cluster, complete sequence.
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McQuaid, J.,
1 (bases 1 to 230000)
TITLE Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL MEDLINE 20510030
PUBMED 11054574
REFERENCE 2 (bases 1 to 230000)
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McQuaid, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Direct Submision
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
source location/Qualifiers
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/complement(<6963..>11517)
/gene="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..>11517))

| | |
|------|---|
| CDS | <p>/gene="KLK1" /product="renal kallikrein" complement(join(6963..7119,7668..7804,7923..8212,9482..9641,11472..11517)) /gene="KLK1" /note="serine protease" /codon_start=1 /product="renal kallikrein" /protein_id="AA033353.1" /db_xref="GI:11244758" /translation="MMFLVLCALSLSGTGAPPIOSRIYGVGNECEQHSPQWQALYHFSYFQCGGILVHROWVLTAAHCISDNYQLGHNLPDENTAOVAVSHSPHGFNMSLEHNRADDEYSHDMLRLSEPADITTAVKVLEPTREEVSGTCLAGMSIEBENFSPPDLOCVLKITLBNDECKRAHOKYTDPLCVHLEGGKDTGVDSGGLMCDGVLCGVTSKRYVECTGTPKPSVANRVLSYKNIETIAENS"</p> |
| mRNA | <p>complement(join(13552..13704,14377..14530,14678..14917,15416..15569,19204..19246)) /product="ACO protease" complement(join(13552..13704,14377..14530,14678..14917,15416..15569,19204..19246)) /note="serine protease" /codon_start=1 /product="ACO protease" /protein_id="AA033354.1" /db_xref="GI:11244759" /translation="MMFLTLFLSLSTAODGKLEGEDECAPHSGQPMQVALYERGRFNGASLISPHWLSAHCOSRFRVRLGHNLRKRDPEQLTTRVTHFRYEARSHRNDMLRLVOPARLNPQVPAVLPTRCPHPEACVSGMGLVPLSSPSPDLTHCANISIIISDSCDSYPGRLTNVWCAEGRGESCEGSGGLVCGGIIQGIIVSGWDVPCDNTKPKGVYKVCYHEWIRETMKRN"</p> |
| gene | <p><42595..>47769 /gene="KLK3" join(42595..42640,43880..44039,45669..45955,46099..46235,47614..47769) /gene="KLK3" /product="prostate specific antigen" join(42595..42640,43880..44039,45669..45955,46099..46235,47614..47769) /gene="KLK3" /note="serine protease" /codon_start=1 /product="prostate specific antigen" /protein_id="AA033355.1" /db_xref="GI:11244760" /translation="MMVPVYFLTLSTWIGAALILSRIVGNECEKHSQPMQVALVSRGAVCGVLVHPQWVLTAAHCIRNKSVALIGRHSLEPHEDTQVQVSHSPHPLYDMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVMDLPTQEPALGTTCYASGMSIEPEEFLTPKVLQCVDLHVISNDYCAQVHPQKTKFPMLCAGRTGKSTCSGDSGGLVCNGVLGITSWGSEPCALPERPSLYTKVHYRKIMDTIANP"</p> |
| mRNA | <p><61139..>66229 /gene="KLK2" join(61139..61184,62391..62550,64142..64428,64542..64678,66074..66229) /gene="KLK2" /product="glutular kallikrein" join(61139..61184,62391..62550,64142..64428,64542..64678,66074..66229) /note="serine protease" /codon_start=1 /product="glutular kallikrein" /protein_id="AA033356.1" /db_xref="GI:11244761" /translation="MMDLVLSIALSVGCTGAVPLIOSRIYGVGNECEKHSQPMQVALVSGHAGCGVLVHPQWVLTAAHCIRNKSQVLTGHNLFEPEDTQVQVSHSPHPLYNMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVGLPTQEPALGTTCYASGMSIEPEEFLRPRLQCVSLHLSNDWCARAYSEKVEPMICAGLMTGKDTCCGDSGGLVCNGVLGITSWGSEPCALPERPAVYTKVHYRKIMDTIANP"</p> |
| CDS | <p>complement(join(94602..94755,96027..96163,96247..96497,96919..97081,98345..98405)) /product="serine protease" complement(join(94602..94755,96027..96163,96247..96497,96919..97081,98345..98405))</p> |
| gene | <p>/codon_start=1 /product="serine protease" /protein_id="AA033357.1" /db_xref="GI:11244762" /translation="MATAGNPMWGLGYLILGVAGSLVSGSCQIINGEDCSPHSPMQALVWENELFCGCVLVHPQWVLTAAHCIRNKSQVLTGHNLFEPEDTQVQVSHSPHPLYNMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVGLPTQEPALGTTCYASGMSIEPEEFLRPRLQCVSLHLSNDWCARAYSEKVEPMICAGLMTGKDTCCGDSGGLVCNGVLGITSWGSEPCALPERPAVYTKVHYRKIMDTIANP"</p> |
| mRNA | <p>complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..1410303)) /gene="KLK5" /product="stratum corneum trypsin-like serine protease" complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..1410303)) /gene="KLK5" /note="synonym: SCCE" /codon_start=1 /product="stratum corneum trypsin-like serine protease" /protein_id="AA033358.1" /db_xref="GI:11244763" /translation="MMVLCALITALLGYTDARSDSSRIINGSDCMHTOPQWALILRPNOLYCGAVLVHPQWVLTAAHCIRNKSQVLTGHNLFEPEDTQVQVSHSPHPLYNMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVGLPTQEPALGTTCYASGMSIEPEEFLRPRLQCVSLHLSNDWCARAYSEKVEPMICAGLMTGKDTCCGDSGGLVCNGVLGITSWGSEPCALPERPAVYTKVHYRKIMDTIANP"</p> |
| CDS | <p>complement(join(146834..146986,149628..149764,151186..151433,155052..155208,155948..155987)) /product="protease M" complement(join(146834..146986,149628..149764,151186..151433,155052..155208,155948..155987)) /note="serine protease; also called neurosin or zyme" /codon_start=1 /product="protease M" /protein_id="AA033359.1" /db_xref="GI:11244764" /translation="MKKLMVYSLTAAAEQRNQLVHGPGDKTSHPYOALYTSCHLTCGGVLHPVLTAAHCIRNKSQVLTGHNLFEPEDTQVQVSHSPHPLYNMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVGLPTQEPALGTTCYASGMSIEPEEFLRPRLQCVSLHLSNDWCARAYSEKVEPMICAGLMTGKDTCCGDSGGLVCNGVLGITSWGSEPCALPERPAVYTKVHYRKIMDTIANP"</p> |
| mRNA | <p>complement(join(165420..165575,167672..167808,168124..168371,169651..169798,170211..170283)) /gene="KLK7" /product="stratum corneum chymotryptic enzyme" complement(join(165420..165575,167672..167808,168124..168371,169651..169798,170211..170283)) /gene="KLK7" /note="synonym: SCCE" /codon_start=1 /product="stratum corneum chymotryptic enzyme" /protein_id="AA033360.1" /db_xref="GI:11244765" /translation="MARSLLPLQLILLSLAETAGEBAQDKITDGACGANGSHPMQVALLSGNOLHCGVLVHPQWVLTAAHCIRNKSQVLTGHNLFEPEDTQVQVSHSPHPLYNMSLLKORFLRPBDDSSHDLMLRLSEBALTDAVKVGLPTQEPALGTTCYASGMSIEPEEFLRPRLQCVSLHLSNDWCARAYSEKVEPMICAGLMTGKDTCCGDSGGLVCNGVLGITSWGSEPCALPERPAVYTKVHYRKIMDTIANP"</p> |
| gene | <p>complement(join(183943..184098,185635..185768,187865..188127,188293..188452,188967..1890361)) /product="neutropin" complement(join(183943..184098,185635..185768,187865..188127,188293..188452,188967..1890361)) /note="serine protease; also called ovaasin" /codon_start=1 /product="neutropin" /protein_id="AA033361.1" /db_xref="GI:11244766"</p> |

/translation="MGRPRRAKTMFLLLIGAMWAGHSAQEDKVLGHCORHSD
PWOALFOGOQLICGVLVGVGMVLLTAACSKPKYLVADSLQNKDREPIVVO
SLPHCYNSDVEDENHNDMLLORLOASLQKXKPISLADHCTOPKOCSTYSGMGV
TSPRNPPTLNCACVKKIPROKCEDAYGOITDGMVCGSSKSGADTCQDGBGFLVC
DGAAGITSMSDPCGRSPKREVTNICYLDMIKKIJSSKG
complement(join<190980..191129,191573..191709,
194324..194589,197048..197204,197370..1974121)
/product="Kallikrein-like 3"
complement(join(190980..191129,191573..191709,
194324..194589,197048..197204,197370..1974121)
/note="serine protease"
/codon_start=1
/product="Kallikrein-like 3"
/protein_id="AA033362.1"
/db_xref="GI:11244767"
/translation="MKILGALSLAGHMDRATIGAEBCRPNOSPMOAGLPHLT
RLFCATLISDWLLTAACRPRYLMLVLEHNLKMKWERBOLFVTFPRPRKND
LSANDHNDIMLIRLPROKLSRAVOPPLSLSCVSPKQCLISGMVAISSKALPV
TLQCANISLENLCHMAYPGHISDMLCAEMEGRGSCQDSDGPIVNCNGLTAGVV
SGAEPSCRPRRAVYTSVCHLDMIOEIMEN"

Query Match 100.0%; Score 1571; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGTATGGGATTTAGAGACAGAAACAAATTAGTCCAGAGATGATTTCTA 60
DB 132548 AGGAGGTATGGGATTTAGAGACAGAAACAAATTAGTCCAGAGATGATTTCTA 132489
QY 61 TTGGAGATGATTTCTGCTTAAAGACATGGCAATCCAGAGATTTGGTTGTCA 120
DB 132488 TTGGAGATGATTTCTGCTTAAAGACATGGCAATCCAGAGATTTGGTTGTCA 132429
QY 121 CAATATATGAGGGGCACTTACGCGCACTAATGATAGTCAAGTGTGCTTCAAC 180
DB 132428 CAATATATGAGGGGCACTTACGCGCACTAATGATAGTCAAGTGTGCTTCAAC 132369
QY 181 ATGCTATATGACACCGGAGGCTTCCACAAACCAATTATCCAGCTTCAGATGCCAC 240
DB 132368 ATGCTATATGACACCGGAGGCTTCCACAAACCAATTATCCAGCTTCAGATGCCAC 132309
QY 241 AGTGGCCAGATGAGGAACCTCTATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 300
DB 132308 AGTGGCCAGATGAGGAACCTCTATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 132249
QY 301 GTATAGATGAGGATTTGTTGAGAGATGGGAGAAAGTGTGTGTCAGTAAGAAATTA 360
DB 132248 GTATAGATGAGGATTTGTTGAGAGATGGGAGAAAGTGTGTGTCAGTAAGAAATTA 132189
QY 361 GGCCTGACAGGCTGAGGGGAGAGTGAAGAAAGGAGGCGGAGAGATACAGATGA 420
DB 132188 GGCCTGACAGGCTGAGGGGAGAGTGAAGAAAGGAGGCGGAGAGATACAGATGA 132129
QY 421 GGGAGACAGGCTGAGAACAGAAAGTGAAGACGAAGTGTGAGAGAGAAAGGTTCA 480
DB 132128 GGGAGACAGGCTGAGAACAGAAAGTGAAGACGAAGTGTGAGAGAGAAAGGTTCA 132069
QY 481 CAGACCCCCCGGAAATGATGTGAGCAACAGAAATCTGAGAGAGAAATGAGGTGAG 540
DB 132068 CAGACCCCCCGGAAATGATGTGAGCAACAGAAATCTGAGAGAGAAATGAGGTGAG 132009
QY 541 AGTGACAAATGGGCTTAAAGTTGAACCTTGAGGCGAGAGATGTGCTCAGCCTGTA 600
DB 132008 AGTGACAAATGGGCTTAAAGTTGAACCTTGAGGCGAGAGATGTGCTCAGCCTGTA 131949
QY 601 ATCCCAACACTTTGAGGCTGAGGTGGCGGAATCACTTGAGGCGAGAGTTGAGACCG 660
DB 131948 ATCCCAACACTTTGAGGCTGAGGTGGCGGAATCACTTGAGGCGAGAGTTGAGACCG 131889
QY 661 CTTGGCAACATGTGGAACCCGCTCTACAAAATAATTAATAAATAATTAACCGGGTGT 720
DB 131888 CTTGGCAACATGTGGAACCCGCTCTACAAAATAATTAATAAATAATTAACCGGGTGT 131829

QY 721 GGTGATGACACCTGTAAGTCAACGTACTTGGAGAGCTGAGGACAGAGATTTGTTGAC 780
DB 131828 GGTGATGACACCTGTAAGTCAACGTACTTGGAGAGCTGAGGACAGAGATTTGTTGAC 131769
QY 781 CCGGAGATGAGGCTGAGTGAAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 840
DB 131768 CCGGAGATGAGGCTGAGTGAAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 131709
QY 841 GAGTAAGATCAGTCAAAAAAAGAGTGAATTTGAGTGAATATTAATAACAT 900
DB 131708 GAGTAAGATCAGTCAAAAAAAGAGTGAATTTGAGTGAATATTAATAACAT 131649
QY 901 CTCCCT 960
DB 131648 CTCCCT 131589
QY 961 CTGTAATTTCAATCTCTGTGTGTCTGTGTTCCATCTGCTTCTCAATCTATGAGGATCTCTG 1020
DB 131588 CTGTAATTTCAATCTCTGTGTGTCTGTGTTCCATCTGCTTCTCAATCTATGAGGATCTCTG 131529
QY 1021 GGTCTCTGATGCT 1080
DB 131528 GGTCTCTGATGCT 131469
QY 1081 TCTCTCTGAGGATGATTTCTGGGGGCTGTGTGTCTGCAATGAGTCTCTCTGAGGAGCTC 1140
DB 131468 TCTCTCTGAGGATGATTTCTGGGGGCTGTGTGTCTGCAATGAGTCTCTCTGAGGAGCTC 131409
QY 1141 GTGTCTTGGGAGATTAACCTTTGTGCTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1200
DB 131408 GTGTCTTGGGAGATTAACCTTTGTGCTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 131349
QY 1201 TGCAAGTTCACCAATGTGATCCAGGAAACCAATCCAGGCAACTCTGATGATCCAGGAA 1260
DB 131348 TGCAAGTTCACCAATGTGATCCAGGAAACCAATCCAGGCAACTCTGATGATCCAGGAA 131289
QY 1261 CTGAGCACACCGGATCCCACTGCTGAGGAGCAGGCTGAGCACTCTTTCAGACCT 1320
DB 131288 CTGAGCACACCGGATCCCACTGCTGAGGAGCAGGCTGAGCACTCTTTCAGACCT 131229
QY 1321 CATTCCTTCCAGAGATTTGAGAAATTTCAATCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1380
DB 131228 CATTCCTTCCAGAGATTTGAGAAATTTCAATCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 131169
QY 1381 GGAATCAGGATCTGCTTCCCAACATTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1440
DB 131168 GGAATCAGGATCTGCTTCCCAACATTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 131109
QY 1441 AACATTTCCAAACCTGTCAGGGGAGGCTTGCCTCAATCTCTGAGGCTGAGGCTGAGGCTGAG 1500
DB 131108 AACATTTCCAAACCTGTCAGGGGAGGCTTGCCTCAATCTCTGAGGCTGAGGCTGAGGCTGAG 131049
QY 1501 TCTCAAGCTCAGGGGCTGATCTCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1560
DB 131048 TCTCAAGCTCAGGGGCTGATCTCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 130989
QY 1561 AAACCTGAGAAG 1571
DB 130988 AAACCTGAGAAG 130978

RESULT 7
AC027602/c
LOCUS
DEFINITION
SEQUENCE, 12 unordered pieces.
ACCESSION
AC027602
VERSION
AC027602.4 GI:11178143
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 217346)
Birnén, B., Linton, L., Nuebaum, C. and Lander, E.
Homo sapiens chromosome 19, clone RP11-795B6
Unpublished
2 (bases 1 to 217346)
Birnén, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choegl, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, D.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Kleith, J., Larocque, K., Lamazeres, R., Landers, T., Lehocsky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
McCarthy, M., McKean, P., McGuirk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisanti, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 217346)
Birnén, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Baetien, V., Beda, F., Boguslavsky, L.,
Boukhalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A.,
Choegl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, D.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,
Lamazeres, R., Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Margulis, N., McCarthy, M., McKean, P., McKernan, K.,
McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K.,
Pierre, N., Pisanti, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 15, 2000 this sequence version replaced gi:11136831.
All repeats were identified using RepeatMasker:
Smit, A. P. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L9166
Center Clone name: 795_B_6
----- Summary Statistics
Sequencing vector: M13; M77815; 31% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 213655 bases at least Q30
Consensus quality: 215056 bases at least Q20

Insert size: 194000; agarose-fp
Insert size: 216246; sum-of-ctrls
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.
NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
8150 8249: contig of 8149 bp in length
8250 9592: contig of 1342 bp in length
9593 9692: gap of 100 bp
9693 10733: contig of 1041 bp in length
10734 10833: gap of 100 bp
10834 13519: contig of 2686 bp in length
13520 13619: gap of 100 bp
13620 17510: contig of 3891 bp in length
17511 17610: gap of 100 bp
17610 24602: contig of 6992 bp in length
24603 24702: gap of 100 bp
24703 35434: contig of 10732 bp in length
35435 35534: gap of 100 bp
35535 124474: contig of 88940 bp in length
124475 124574: gap of 100 bp
124575 134664: contig of 10090 bp in length
134665 134764: gap of 100 bp
134765 162344: contig of 27579 bp in length
162344 162444: gap of 100 bp
162444 208917: contig of 46474 bp in length
208918 209018: gap of 100 bp
209018 217346: contig of 8329 bp in length.
Location/Qualifiers
1. 217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_11b="RP11-795B6"
1. 8149
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
8250..9592
/note="assembly_fragment"
9693..10733
/note="assembly_fragment"
10834..13519
/note="assembly_fragment"
13620..17510
/note="assembly_fragment"
17611..24602
/note="assembly_fragment"
24703..35434
/note="assembly_fragment"
35535..124474
/note="assembly_fragment"
124575..134664
/note="assembly_fragment"
134765..162343
/note="assembly_fragment"
162444..208917
/note="assembly_fragment"
209018..217346
/note="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN

Query Match 99.0%; Score 1555.2; DB 2; Length 217346;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1568; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1 AGGAGATATGGGGAATTTGAAGACAGAAACAAATTTAGTCCAGCGAATGGATTCTA 60
194423 AGGAGATATGGGGAATTTGAAGACAGAAACAAATTTAGTCCAGCGAATGGATTCTA 194364

61 TTGGAGATATCTGCCCCCTAGAAAGACATGGCAATACAGAGAGACATTTTGGTTGTA 120
194363 TTGGAGATATCTGCCCCCTAGAAAGACATGGCAATACAGAGAGACATTTTGGTTGTA 194304

121 CAATATATGAGGGGCAATTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 180
194303 CAATATATGAGGGGCAATTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 194244

181 ATGCTATGATGACACAGGAGGCTTCAACAACAACATTTATTCAGCTTCAAGTCCAC 240
194243 ATGCTATGATGACACAGGAGGCTTCAACAACAACATTTATTCAGCTTCAAGTCCAC 194184

241 AGTCCAGATGAGAGACCTCATTCAGGGGCTGAGAACCGATTTTGGAGAGGAG 300
194183 AGTCCAGATGAGAGACCTCATTCAGGGGCTGAGAACCGATTTTGGAGAGGAG 194124

301 GTATAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 360
194123 GTATAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 194064

361 GGCCTGACACAGGCTGAGAGGAGAGATGAGAGAGAGAGAGAGAGAGAGATACAGTAA 420
194063 GGCCTGACACAGGCTGAGAGGAGAGATGAGAGAGAGAGAGAGAGAGATACAGTAA 194004

421 GGGAGACAGGCTGAGAGAGAGATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 480
194003 GGGAGACAGGCTGAGAGAGAGATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 193944

481 CAGACCCCCCAAAATGATGTGACAAACAGGAATCTGAGAGAGAGAGATGAGATGAG 540
193943 CAGACCCCCCAAAATGATGTGACAAACAGGAATCTGAGAGAGAGATGAGATGAGATGAG 193884

541 AGTGAACAATGGGCTTTAAAGTTGAACCTTGAGGCTGAGGCTCAGCCTGTA 600
193883 AGTGAACAATGGGCTTTAAAGTTGAACCTTGAGGCTGAGGCTCAGCCTGTA 193824

601 ATCCCAACCTTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 660
193823 ATCCCAACCTTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 193764

661 CCTGGCCAAATGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 720
193763 CCTGGCCAAATGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 193704

721 GGTGATGACACCTGTAATGACACTTCTGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 780
193703 GGTGATGACACCTGTAATGACACTTCTGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 193644

781 CCGGAGATGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 840
193643 CCGGAGATGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 193584

841 GAGTAAAGATCATCTC-AAAAAAAAAAAAAGCTGATTTGAGTGAATATTAATACAT 899
193583 GAGTAAAGATCATCTC-AAAAAAAAAAAAAGCTGATTTGAGTGAATATTAATACAT 193524

900 TCTCCCTCTCTCTCTCTTTGGCTGTGTCTCATCTGTCTTTTTCGCAATTTCTTCATC 959
193523 TCTCCCTCTCTCTCTCTTTGGCTGTGTCTCATCTGTCTTTTTCGCAATTTCTTCATC 193464

960 TCTGATCTTTCATCTCTGTGTGTCTGTCCATGCTTTCATCTTTCATCTTTCATCTTTCAT 1019
193463 TCTGATCTTTCATCTCTGTGTGTCTGTCCATGCTTTCATCTTTCATCTTTCATCTTTCAT 193404

1020 GGGTCTCATGATCTCTTTCGCGCACTTTCGACATCTTTCGCTCTCATGCTCCCT 1079
193403 GGGTCTCATGATCTCTTTCGCGCACTTTCGACATCTTTCGCTCTCATGCTCCCT 193344

1080 TTCTCTCTGACAGGATTTCTGGGGGCTGTGTCTGCAATGCTCTCCGACAGGACT 1139
193343 TTCTCTCTGACAGGATTTCTGGGGGCTGTGTCTGCAATGCTCTCCGACAGGACT 193284

1140 CGTGTCTGGGAGATTAACCTTGTGCGCGGCAACAGACCGGCTGTCTACAGAACT 1199
193283 CGTGTCTGGGAGATTAACCTTGTGCGCGGCAACAGACCGGCTGTCTACAGAACT 193224

1200 CTGCAAGTTCAACAGATGATTCAGAGAAACATCCAGGCAATCTCTGATCATCCAG 1259
193223 CTGCAAGTTCAACAGATGATTCAGAGAAACATCCAGGCAATCTCTGATCATCCAG 193164

1260 ACTGACACACCGGATTCCTCCACCTGTGACAGGACAGCCTGAAATCTCTTCAAGCC 1319
193163 ACTGACACACCGGATTCCTCCACCTGTGACAGGACAGCCTGAAATCTCTTCAAGCC 193104

1320 TCATTCCTTCCAGAGATTTGAGATGTTCACTCTGCAAGCCCTGACCCCATGCTCC 1379
193103 TCATTCCTTCCAGAGATTTGAGATGTTCACTCTGCAAGCCCTGACCCCATGCTCC 193044

1380 TGGACTCAGGCTCTGCTTCCCTCCACATTTGAGGCTGACGCTCTCTAGTTAACTG 1439
193043 TGGACTCAGGCTCTGCTTCCCTCCACATTTGAGGCTGACGCTCTCTAGTTAACTG 192984

1440 GAACAATTTCCAAATCTGTCAGAGGCGGGGCTGTGCTCAATCTTCTGAGGCTTTC 1499
192983 GAACAATTTCCAAATCTGTCAGAGGCGGGGCTGTGCTCAATCTTCTGAGGCTTTC 192924

1500 ATCTCAAGCTCAGGAGCCATCTCTCTGAGCTGACCAATTTAGTCCAGAAA 1559
192923 ATCTCAAGCTCAGGAGCCATCTCTCTGAGCTGACCAATTTAGTCCAGAAA 192864

1560 TAACTGAGAG 1571
192863 TAACTGAGAG 192852

RESULT 8
AC130782/c
LOCUS
DEFINITION
ORDERED pieces.
AC130782.2 GI:25167101
VERSION
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
1 (bases 1 to 200792)
Akhter,N., Antonellis,R., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakeley,A.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Latic,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.U., Maduro,V.B.,
Maruligan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddik-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stancir,D., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 200792)
Green,E.D.
Direct Submission
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovecrest Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
Green,E.D.
REFERENCE
AUTHORS

TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
COMMENT On Nov 22, 2002 this sequence version replaced gi:22216452.

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dhx
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data, from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

1 28306: contig of 28306 bp in length
28307 28406: gap of unknown length
28407 37856: contig of 9450 bp in length
37857 37956: gap of unknown length
37957 73522: contig of 35566 bp in length
73523 73622: gap of unknown length
73623 83567: gap of 9945 bp in length
83568 83667: gap of unknown length
83668 88817: contig of 5150 bp in length
88818 88917: gap of unknown length
88918 125611: contig of 36694 bp in length
125612 125711: gap of unknown length
125712 159879: contig of 34168 bp in length
159880 159979: gap of unknown length
159980 174698: contig of 14719 bp in length
174699 174798: gap of unknown length
174799 186382: contig of 11584 bp in length
186383 186482: gap of unknown length
186483 193344: contig of 6862 bp in length
193345 193444: gap of unknown length
193445 193633: contig of 5919 bp in length
193634 199463: gap of unknown length
199464 200792: contig of 1329 bp in length.

FEATURES
Source
1. .200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"

/db xref="taxon:9598"
/clone="CH251-355A20"
/clone_11b="CH251"
1. .28306
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
28407..37856
/note="assembly_fragment"
37957..73522
/note="assembly_fragment"
73623..83567
/note="assembly_fragment"
83668..88817
/note="assembly_fragment"
88918..125611
/note="assembly_fragment"
125712..159879
/note="assembly_fragment"
159980..174698
/note="assembly_fragment"
174799..186382
/note="assembly_fragment"
186483..193344
/note="assembly_fragment"
193445..193633
/note="assembly_fragment"
199464..200792
/note="assembly_fragment"
clone_end:17
vector_side:right"

ORIGIN

Query Match 94.3%; Score 1482; DB 2; Length 200792;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1542; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

1 AGGAGAGTATGGGAGTTTGAAGACAGGAACCAATTTGTCACAGCAATGATTTCTA 60
69290 AGGAGAGTATGGGAGTTTGAAGACAGGAACCAATTTGTCACAGCAATGATTTCTA 69231
QY TTGGAGTAGTATTCGCTCCCTAGAGACACTGCGCAATACAGAGACATTTTGGTTGTC 120
69230 TTGGAGTAGTATTCGCTCCCTAGAGACACTGCGCAATACAGAGACATTTTGGTTGTC 69171
Db
QY 121 CAATATATGAGAGGCGCATTAATGCACTAATGATAGATGCCAATGCTGTTCAAC 180
69170 CAATATATGAGAGGCGCATTAATGCACTAATGATAGATGCCAATGCTGTTCAAC 69111
Db
QY 181 ATGCTATGATGACACAGGCGCTTCACCAACCAATTTACAGCTTCAGATGCCAC 240
69110 ATGCTATGATGACACAGGCGCTTCACCAACCAATTTACAGCTTCAGATGCCAC 69051
Db
QY 241 AGTCCAGATGAGAGACCTCATCCAGAGGCTGGAACCGATTTTGGCAGAGAGAG 300
69050 AGTCCAGATGAGAGACCTCATCCAGAGGCTGGAACCGATTTTGGCAGAGAGAG 68991
Db
QY 301 GATATAGATGAGTGGTGGTGGAGATGGGAGAGAGTGTGTCCAGTAAGAAATTA 360
68990 GATATAGATGAGTGGTGGTGGAGATGGGAGAGAGTGTGTCCAGTAAGAAATTA 68931
Db
QY 361 GGCCTGCAACAGCTGAGAGGAGAGTGAAGAGAGAGAGGCGGAGATACAGATGA 420
68930 GGCCTGCAACAGCTGAGAGGAGAGTGAAGAGAGAGAGGCGGAGATACAGATGA 68871
Db
QY 421 GGGAGACAGCTGGAACAGAAAGTGAAGAGAGATTCAGAGTGTGAGAGAGAGGCTCA 480
68870 GGGAGACAGCTGGAACAGAAAGTGAAGAGAGATTCAGAGTGTGAGAGAGAGGCTCA 68811
Db
QY 481 GAGACCCCGCAAAATGATGTGTGCAACAGAAATCTGGAAGAGAGATGATGTGAG 540
68810 GAGACCCCGCAAAATGATGTGT-GACAAACAGAAATCTGGAAGAGAGATGATGTGAG 68753

QY 541 AGTGA CAAATGGGGTCTAAAGTTGAACTTGAGGCGCAGGCAATGATGCTCAAGCCTGTA 600
AX331407/c
Db 68752 AGTGA CAAATGGGGTCTGAAAGTTGAACTTGAGGCGCAGGCAATGATGCTCAAGCCTGTA 68693
QY 601 ATCCCAACCTTTGAGGGCTGAGAGTGGGGAATCACTTGAGGCGCAGGAGTTGAGACCGA 660
Db 68692 ATCCCAATCTTTGAGGGCTGAGAGTGGGGAATCACTTGAGGCGCAGGAGTTGAGACCGA 68633
QY 661 CTTGGCCCAATGATGTAACCCCGTCTAC--AAAAAAATACAAAAATTTAGCCGGGNG 719
Db 68632 CTTGGCCCAATGATGTAACCCCGTCTACAAAAAAATACAAAAATTTAGCCGGGNG 68573
QY 720 TGGTATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGCGAGAGATTGCTTGA 779
Db 68572 TGGTATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGCGAGAGATTGCTTGA 68513
QY 780 CCGGGAGATGAGAGCTGAGAGTGAAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 839
Db 68512 CCGGGAGATGAGAGCTGAGAGTGAAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 68453
QY 840 AGATGAAGCTCATCTCAAAAAAAGCTGATTTGAGTGAATATTATTAACAT 899
Db 68452 AGATGAAGCTCATCTC--AAAAAAGCTGATTTGAGTGAATATTATTAACAT 68395
QY 900 TCTCCCT 959
Db 68394 TCTCCCT 68335
QY 960 TCTGTAATTTTCAATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1019
Db 68334 TCTGTAATTTTCAATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 68275
QY 1020 GGGTCTCTCATGCTCTCTCTCTGCGCACTTGGCAACATCTCTGCTCTCATGCCCCCT 1079
Db 68274 GGGTCTCTCATGCTCTCTCTCTGCGCACTTGGCAACATCTCTGCTCTCATGCCCCCT 68215
QY 1080 TTCTCTCTGAGGATGATTTCTGGGGGCGCTGATGCTGCAATGATGCTGCAAGGACT 1139
Db 68214 TTCTCTCTGAGGATGATTTCTGGGGGCGCTGATGCTGCAATGATGCTGCAAGGACT 68155
QY 1140 CGTGTCTGGGAGATTAACCTTTGTCGCGGCGCAACAGACCGGGTGTCTACACGACT 1199
Db 68154 CGTGTCTGGGAGATTAACCTTTGTCGCGGCGCAACAGACCGGGTGTCTACACGACT 68095
QY 1200 CTGCAAGTTCAACAGTGTATCAGAGAAACATCCAGGCACTCTGAGTCAATCCAG 1259
Db 68094 CTGCAAGTTCAACAGTGTATCAGAGAAACATCCAGGCACTCTGAGTCAATCCAG 68035
QY 1260 ACTGAGCACCGGCAATCCCACTGCTGAGGAGACGCTGACACTCTCTTCAAGACC 1319
Db 68034 ACTGAGCACCGGCAATCCCACTGCTGAGGAGACGCTGACACTCTCTTCAAGACC 67975
QY 1320 TCAATTCCTTCCAGAGATGTTGAGATGTTCACTCTCAAGCCCTGACCCCATGCTTC 1379
Db 67974 TCAATTCCTTCCAGAGATGTTGAGATGTTCACTCTCAAGCCCTGACCCCATGCTTC 67915
QY 1380 TGGATCTAGAGGCTGCTGCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTG 1439
Db 67914 TGGATCTAGAGGCTGCTGCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTG 67855
QY 1440 GAAACAATTTCAAACTGTCAAGGCGGGGTTGCGTCAATCTCTCCCTGAGGCACTTTC 1499
Db 67854 GAAACAATTTCAAACTGTCAAGGCGGGGTTGCGTCAATCTCTCCCTGAGGCACTTTC 67795
QY 1500 ATCTCAAGCTCAAGGCGCCATCTCTCTCTGCAAGCTCTGACCAATTTAGTCCAGAAA 1559
Db 67794 ATCTCAAGCTCAAGGCGCCATCTCTCTCTGCAAGCTCTGACCAATTTAGTCCAGAAA 67735
QY 1560 TAAAGTGAAG 1571
Db 67734 TAAAGTGAAG 67723

RESULT 9
AX331407/c
LOCUS AX331407 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 1916 from Patent WO0194629.
ACCESSION AX331407
VERSION AX331407.1 GI:18122041
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 1916 13-DEC-2001;
JOURNAL Avalon Pharmaceuticals (US)
FEATURES
source 1..586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 3,8e-11;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGATGATTTCTGGGGGCGCTGTGTGTCGCAATGAGCTCTCTGAGGAGCTGTGTCTG 1149
Db 501 CAGGATGATTTCTGGGGGCGCTGTGTGTCGCAATGAGCTCTCTGAGGAGCTGTGTCTG 442
QY 1150 GGAATTTACCTTTGCCCCCGCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db 441 GGAATTTACCTTTGCCCCCGCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGAATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACA 1269
Db 381 ACCAAGTGAATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACA 322
QY 1270 CCGGATCCCACTGTGTCAGAGGACAGCCCTGACACTCTTTCAGACCTTCATTCCTTC 1329
Db 321 CCGGATCCCACTGTGTCAGAGGACAGCCCTGACACTCTTTCAGACCTTCATTCCTTC 262
QY 1330 CCAAGATGTTGAGATGTTCACTCTCAAGCCCTGACCCCATGCTCTCTGAGTCAAG 1389
Db 261 CCAAGATGTTGAGATGTTCACTCTCAAGCCCTGACCCCATGCTCTCTGAGTCAAG 202
QY 1390 GTCTGCTTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTGAGGCACTTCAAGC 1449
Db 201 GTCTGCTTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTGAGGCACTTCAAGC 142
QY 1450 CAAAATCTGTCAGAGGCGGGGTTGCGTCAATCTCTCTGAGGCACTTCAAGC 1509
Db 141 CAAAATCTGTCAGAGGCGGGGTTGCGTCAATCTCTCTGAGGCACTTCAAGC 82
QY 1510 TCAAGGCGCCATCTCTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATTAATGAGA 1569
Db 81 TCAAGGCGCCATCTCTCTCTGCAAGTCTGACCCAAATTTAGTCCAGAAATTAATGAGA 22
QY 1570 AG 1571
Db 21 AG 20
RESULT 10
AX331824/c
LOCUS AX331824 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2333 from Patent WO0194629.
ACCESSION AX331824
VERSION AX331824.1 GI:18122458
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo; 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 2333 13-DEC-2001;
FEATURES Location/Qualifiers
 source 1..586
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 30.7%; Score 482; DB 6; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGGCTCCCTGAGGACTCGTCTCG 1149
 |||||
 DB 501 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGGCTCCCTGAGGACTCGTCTCG 442
 |||||

QY 1150 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
 |||||
 DB 441 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382
 |||||

QY 1210 ACCAAGTGAATCCAGGAACCATCCAGGCACTCCGATGATCCAGGACTCAGACAC 1269
 |||||
 DB 381 ACCAAGTGAATCCAGGAACCATCCAGGCACTCCGATGATCCAGGACTCAGACAC 322
 |||||

QY 1270 CCGGATCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329
 |||||
 DB 321 CCGGATCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262
 |||||

QY 1330 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 1389
 |||||
 DB 261 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 202
 |||||

QY 1390 GTCTGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449
 |||||
 DB 201 GTCTGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
 |||||

QY 1450 CAAAATCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
 |||||
 DB 141 CAAAATCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 82
 |||||

QY 1510 TCAGGAGCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1569
 |||||
 DB 81 TCAGGAGCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 22
 |||||

QY 1570 AG 1571
 |||||
 DB 21 AG 20
 |||||

RESULT 11
 G41906/c 586 bp DNA 1linear STS 30-SEP-1998
LOCUS SHGC-56840 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G41906
ACCESSION G41906.1 GI:3668239
VERSION STS.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS Myers, R.M.
TITLE Human STS (1998)

JOURNAL Unpublished (1998)
COMMENT
 Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA.
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@hgc.stanford.edu
 Primer A: AGAGACAGCTCAGCCCAAT
 Primer B: GCCAATCTGATCATCTCC
 STS size: 188
 PCR Profile:
 Initial incubation: 95 degrees C for 10 minutes
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9700

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

STSS
 primer bind 167.186
 primer_bind complement(335.354)

ORIGIN
 Query Match 30.7%; Score 482; DB 11; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGGCTCCCTGAGGACTCGTCTCG 1149
 |||||
 DB 501 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGGCTCCCTGAGGACTCGTCTCG 442
 |||||

QY 1150 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
 |||||
 DB 441 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382
 |||||

QY 1210 ACCAAGTGAATCCAGGAACCATCCAGGCACTCCGATGATCCAGGACTCAGACAC 1269
 |||||
 DB 381 ACCAAGTGAATCCAGGAACCATCCAGGCACTCCGATGATCCAGGACTCAGACAC 322
 |||||

QY 1270 CCGGATCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329
 |||||
 DB 321 CCGGATCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262
 |||||

QY 1330 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 1389
 |||||
 DB 261 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 202
 |||||

QY 1390 GTCTGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449
 |||||
 DB 201 GTCTGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
 |||||

QY 1450 CAAAATGTCAGAGGGGGGTTGCGTCAATCTCCCTGAGGACATTTTCATCTCAAG 1509
DB 141 CAAAATGTCAGAGGGGGGTTGCGTCAATCTCCCTGAGGACATTTTCATCTCAAG 82
QY 1510 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTAGTCCAGAAATAAAGTGA 1569
DB 81 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTAGTCCAGAAATAAAGTGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 12
AX429955/c 735 bp DNA linear PAT 21-JUN-2002
LOCUS AX429955
DEFINITION Sequence 94 from Patent WO0198339.
ACCESSION AX429955
VERSION AX429955.1 GI:21541119
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0198339-A 94 27-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..735
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 3.9e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCGTCTCG 1149
DB 500 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCGTCTCG 441
QY 1150 GGAGATTACCTTTGTCCTGGCCCAACAGACGGGTGTCTACGAACTCTGCAATTTC 1209
DB 440 GGAGATTACCTTTGTCCTGGCCCAACAGACGGGTGTCTACGAACTCTGCAATTTC 381
QY 1210 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 1269
DB 380 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 321
QY 1270 CCGGATCCCACTGCTGACAGGACAGCCCTGACACTCTCTTACAGACCTCATCTTC 1329
DB 320 CCGGATCCCACTGCTGACAGGACAGCCCTGACACTCTCTTACAGACCTCATCTTC 261
QY 1330 CCAGAGATTGAGAAATTTCAATCTCTCAGCCCTGACCCCAATGTTCTCTGAGTCA 1389
DB 260 CCAGAGATTGAGAAATTTCAATCTCTCAGCCCTGACCCCAATGTTCTCTGAGTCA 201
QY 1390 GTCTGTTCCCAATTTGGGGGTCAGCCGTCCTCTAGTTGAACCTGAGGAAATTTTC 1449
DB 200 GTCTGTTCCCAATTTGGGGGTCAGCCGTCCTCTAGTTGAACCTGAGGAAATTTTC 141
QY 1450 CAAAATGTCAGAGGGGGGTTGCGTCAATCTCCCTGAGGACATTTTCATCTCAAG 1509
DB 140 CAAAATGTCAGAGGGGGGTTGCGTCAATCTCCCTGAGGACATTTTCATCTCAAG 81
QY 1510 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTAGTCCAGAAATAAAGTGA 1569
DB 80 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTAGTCCAGAAATAAAGTGA 21

QY 1570 AG 1571
DB 20 AG 19

RESULT 13
BD139877/c 738 bp DNA linear PAT 18-SEP-2002
LOCUS BD139877
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and methods for their use.
ACCESSION BD139877
VERSION BD139877.1 GI:23234822
KEYWORDS JP 2002507387-A/94.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Reed,S.G. and Xu,J.
TITLE Compounds for immunotherapy and diagnosis of breast cancer and methods for their use
JOURNAL Patent: JP 2002507387-A 94 12-MAR-2002;
CORIXA CORP
COMMENT OS Homo sapiens (human)
PN JP 2002507387-A/94
PD 12-MAR-2002
PF 22-DEC-1998 JP 2000526543
PR 24-DEC-1997 US 08/998253, 24-DEC-1997 US 08/998255 PR
17-JUL-1998 US 09/118627, 17-JUL-1998 US 09/118554 PI STEVEN
G REED, JIANGCHUN XU
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
PC C07K14/47,
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/PC
53,
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00
CC Compounds for immunotherapy and diagnosis
of breast cancer and
CC their use methods for
CC Key
FH source 1..738
FT Location/Qualifiers
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 3.9e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCGTCTCG 1149
DB 503 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCGTCTCG 444
QY 1150 GGAGATTACCTTTGTCCTGGCCCAACAGACGGGTGTCTACGAACTCTGCAATTTC 1209
DB 443 GGAGATTACCTTTGTCCTGGCCCAACAGACGGGTGTCTACGAACTCTGCAATTTC 384
QY 1210 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 1269
DB 383 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 324
QY 1270 CCGGATCCCACTGCTGACAGGACAGCCCTGACACTCTCTTACAGACCTCATCTTC 1329
DB 323 CCGGATCCCACTGCTGACAGGACAGCCCTGACACTCTCTTACAGACCTCATCTTC 264
QY 1330 CCAGAGATTGAGAAATTTCAATCTCTCAGCCCTGACCCCAATGTTCTCTGAGTCA 1389

Db 263 CCAGAGATGTGAAGATGTTGTCATCTCTCCAGCCCTGACCCCAATGTCCTGAGACTCAG 204
Qy 1390 GTCTGCTTCCCAATGTTGGAGTGGAGTGTCTCTCTAGTTGAACCTTGGAAACAATTTTC 1449
Db 203 GTCTGCTTCCCAATGTTGGAGTGGAGTGTCTCTCTAGTTGAACCTTGGAAACAATTTTC 144
Qy 1450 CAAAATGTCGAGGGGGGGGGTTCGCTCAATCTCCCTGGGCACTTATCTCAAGC 1509
Db 143 CAAAATGTCGAGGGGGGGGGTTCGCTCAATCTCCCTGGGCACTTATCTCAAGC 84
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 83 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 24
Qy 1570 AG 1571
Db 23 AG 22

RESULT 14
CQ720600 1143 bp DNA linear PAT 03-FEB-2004
LOCUS CQ720600
DEFINITION Sequence 6534 from Patent WO02068579.
ACCESSION CQ720600
VERSION CQ720600.1 GI:42281457
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 6534 06-SEP-2002;
PE Corporation (NY) (US)
LOCATION/Qualifiers
source 1. 1143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 4e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGGATCTTGGGGGGCTGTGTCGTAATGGCTCCCTGAGAGGACTCGTCTCG 1149
Db 657 CAGGGGATCTTGGGGGGCTGTGTCGTAATGGCTCCCTGAGAGGACTCGTCTCG 716
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 1209
Db 717 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 776
Qy 1210 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTCTAGTCAATCCAGGACTCAGACA 1269
Db 777 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTCTAGTCAATCCAGGACTCAGACA 836
Qy 1270 CCGGCAATCCCACTGCTGTGAGAGGACGACCTTCACTCTTTCAGACCTTCACTTCCTTC 1329
Db 837 CCGGCAATCCCACTGCTGTGAGAGGACGACCTTCACTCTTTCAGACCTTCACTTCCTTC 896
Qy 1330 CCAGAGATGTGAAGATGTTGATCTCTCAGCCCTGACCCCAATGTCCTGAGACTCAG 1389
Db 897 CCAGAGATGTGAAGATGTTGATCTCTCAGCCCTGACCCCAATGTCCTGAGACTCAG 956
Qy 1390 GTCTGCTTCCCAATGTTGGAGTGGAGTGTCTCTCTAGTTGAACCTTGGAAACAATTTTC 1449
Db 957 GTCTGCTTCCCAATGTTGGAGTGGAGTGTCTCTCTAGTTGAACCTTGGAAACAATTTTC 1016
Qy 1450 CAAAATGTCGAGGGGGGGGGTTCGCTCAATCTCCCTGGGCACTTATCTCAAGC 1509

Db 1017 CAAAATGTCGAGGGGGGGGGTTCGCTCAATCTCCCTGGGCACTTATCTCAAGC 1076
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 1077 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1136
Qy 1570 AG 1571
Db 1137 AG 1138

RESULT 15
AY461805 1301 bp mRNA linear PRI 03-DEC-2003
LOCUS AY461805
DEFINITION Homo sapiens kallikrein 5 isoform 3 preproprotein (KLK5) mRNA,
complete cds; alternatively spliced.
ACCESSION AY461805
VERSION AY461805.1 GI:38564772
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Michael, I.P., Kurlender, L., Du, D.C. and Diamond, R.P.
Cloning of new splice variants of the human Kallikrein gene 5
Unpublished
REFERENCE 2
Michael, I.P., Kurlender, L., Du, D.C. and Diamond, R.P.
Direct Submission
Submitted (10-NOV-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
LOCATION/Qualifiers
source 1. 1301
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. 1301
/gene="KLK5"
12..626
/gene="KLK5"
/note="alternatively spliced"
/codon_start=1
/product="kallikrein 5 isoform 3 preproprotein"
/protein_id="AAR23814.1"
/db_xref="GI:38564773"
/translation="MATATPPMMWVLCALITLALGTVBHVLANNDVSCDHPSTVPS
GSNODIGAGAGDARDSDSSRIINGSDCDMTOPQALILRPNDLYCGAVLVHFW
LITAAHCRKVRVRLGHSLSPIVSGQMGQVYSIHPGVSHGSHNDMLIKLN
RIRPTKDVPRINVSHPASGATKCLVSGMTTKSPQGRCPSS"

ORIGIN
Query Match 30.7%; Score 482; DB 9; Length 1301;
Best Local Similarity 100.0%; Pred. No. 4.1e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGGATCTTGGGGGGCTGTGTCGTAATGGCTCCCTGAGAGGACTCGTCTCG 1149
Db 820 CAGGGGATCTTGGGGGGCTGTGTCGTAATGGCTCCCTGAGAGGACTCGTCTCG 879
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 1209
Db 880 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 939
Qy 1210 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTCTAGTCAATCCAGGACTCAGACA 1269
Db 940 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTCTAGTCAATCCAGGACTCAGACA 999
Qy 1270 CCGGCAATCCCACTGCTGTGAGAGGACGACCTTCACTCTTTCAGACCTTCACTTCCTTC 1329

```

Db      1000  |||||
QY      1330  |||||
Db      1060  |||||
QY      1390  |||||
Db      1120  |||||
QY      1450  |||||
Db      1180  |||||
QY      1510  |||||
Db      1240  |||||
QY      1570  |||||
Db      1300  |||||

```

Search completed: February 26, 2005, 06:24:09
 Job time : 6702.88 secs

| | | | | | | | | |
|----|--------|-------|--------|---|----------|---|----------|-----------|
| 1 | 5000 | 100.0 | 11570 | 6 | CQ788219 | ✓ | CQ788219 | Sequence |
| 2 | 5000 | 100.0 | 11570 | 6 | CQ874885 | ✓ | CQ874885 | Sequence |
| 3 | 5000 | 100.0 | 11570 | 6 | CQ874961 | ✓ | CQ874961 | Sequence |
| 4 | 5000 | 100.0 | 11570 | 9 | AF135028 | ✓ | AF135028 | Homo sapi |
| 5 | 4989 | 99.8 | 107487 | 9 | AF135143 | ✓ | AF135143 | Homo sapi |
| 6 | 4989 | 99.8 | 230000 | 9 | AF234357 | ✓ | AF234357 | Homo sapi |
| 7 | 4243.4 | 84.9 | 217346 | 2 | AC027602 | ✓ | AC027602 | Homo sapi |
| 8 | 4106.4 | 82.1 | 200792 | 2 | AC130782 | ✓ | AC130782 | Homo sapi |
| 9 | 1035.2 | 20.7 | 176647 | 2 | AC130188 | ✓ | AC130188 | Pan trogl |
| 10 | 451.6 | 9.2 | 32799 | 9 | AC137551 | ✓ | AC137551 | Homo sapi |
| 11 | 457.2 | 9.1 | 190603 | 2 | AC149102 | ✓ | AC149102 | Papio anu |
| 12 | 456.6 | 9.1 | 141923 | 2 | AC109997 | ✓ | AC109997 | Homo sapi |
| 13 | 455 | 9.1 | 138849 | 9 | HS95C20 | ✓ | 297181 | Human DNA |
| 14 | 455 | 9.1 | 171697 | 9 | AC114876 | ✓ | AC114876 | Homo sapi |
| 15 | 455 | 9.1 | 227137 | 9 | AC098461 | ✓ | AC098461 | Homo sapi |
| 16 | 452.4 | 9.0 | 119147 | 9 | AC098873 | ✓ | AC098873 | Homo sapi |
| 17 | 451.6 | 9.0 | 210508 | 2 | AC135612 | ✓ | AC135612 | Pan trogl |
| 18 | 450.6 | 9.0 | 108893 | 9 | AC079169 | ✓ | AC079169 | Homo sapi |
| 19 | 450 | 9.0 | 100208 | 9 | AC135468 | ✓ | AC135468 | Homo sapi |

| | | | | | | |
|------|-------|-----|--------|---|-----------|---------------------|
| 20 | 450 | 9.0 | 182569 | 9 | AC0932500 | AC0925050 Homo sapi |
| C 21 | 449.8 | 9.0 | 119555 | 9 | EX004827 | EX004827 Human DNA |
| C 22 | 449.6 | 9.0 | 147122 | 9 | AC090109 | AC090109 Homo sapi |
| 23 | 449 | 9.0 | 183690 | 9 | AC091180 | AC091180 Homo sapi |
| C 24 | 447.4 | 8.9 | 99876 | 9 | HSB090108 | AL078461 Human DNA |
| C 25 | 444 | 8.9 | 246240 | 6 | AR036572 | AR036572 Sequence |
| C 26 | 444 | 8.9 | 246240 | 6 | AR036573 | AR036573 Sequence |
| C 27 | 444 | 8.9 | 246240 | 6 | AR036574 | AR036574 Sequence |
| C 28 | 444 | 8.9 | 246282 | 9 | HSU91338 | U91338 Human herred |
| C 29 | 443.4 | 8.9 | 106277 | 9 | AC068792 | AC068792 Homo sapi |
| C 30 | 442.4 | 8.8 | 101269 | 9 | AC005300 | AC005300 Homo sapi |
| C 31 | 442.4 | 8.8 | 193159 | 9 | AC006946 | AC006946 Homo sapi |
| C 32 | 442.4 | 8.8 | 196662 | 9 | AP001107 | AP001107 Homo sapi |
| C 33 | 442.4 | 8.8 | 198888 | 9 | AC087728 | AC087728 Homo sapi |
| C 34 | 442 | 8.8 | 109907 | 9 | AC082942 | AC082942 Homo sapi |
| C 35 | 442 | 8.8 | 263876 | 9 | AC019233 | AC019233 Homo sapi |
| C 36 | 441.4 | 8.8 | 158003 | 9 | AC022795 | AC022795 Homo sapi |
| C 37 | 440.4 | 8.8 | 81278 | 9 | AC091062 | AC091062 Homo sapi |
| C 38 | 439.6 | 8.8 | 173058 | 9 | AC104564 | AC104564 Homo sapi |
| C 39 | 439.4 | 8.8 | 50054 | 9 | AC093691 | AC093691 Homo sapi |
| C 40 | 439 | 8.8 | 45896 | 9 | HSB89P11 | AL050312 Human DNA |
| C 41 | 439 | 8.8 | 135640 | 9 | AC036178 | AC036178 Homo sapi |
| C 42 | 438 | 8.8 | 90007 | 9 | AL139234 | AL139234 Human DNA |
| C 43 | 438 | 8.8 | 95930 | 9 | AC106033 | AC106033 Homo sapi |
| C 44 | 438 | 8.8 | 166686 | 9 | AC103895 | AC103895 Papio anu |
| C 45 | 437.6 | 8.8 | 129711 | 9 | AC110768 | AC110768 Homo sapi |

ALIGNMENTS

[illegible]

QY 241 ACACAATGAGACACAGAGGTGTAAAGAAAAGAGATTAA CAGATCCAGATACACCG 300
DB 241 ACACAATGAGACACAGAGGTGTAAAGAAAAGAGATTAA CAGATCCAGATACACCG 300
QY 301 AAAGGGGCGAAGACAGATTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTTT 360
DB 301 AAAGGGGCGAAGACAGATTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTTT 360
QY 361 TTTTTTTTTTTTTTTTGAACGAGTCTCGCTGTGTGCCCAAGGCTGAGAGCGAGCGG 420
DB 361 TTTTTTTTTTTTTTTTGAACGAGTCTCGCTGTGTGCCCAAGGCTGAGAGCGAGCGG 420
QY 421 GATCTGGGCTCACTGCAAGCTCCGCTCCGGGTTACAGGCAATTCCTGCTCAGGCTC 480
DB 421 GATCTGGGCTCACTGCAAGCTCCGCTCCGGGTTACAGGCAATTCCTGCTCAGGCTC 480
QY 481 CCAAGTAGCTGGGACTACAGGCGCCGCACTACGCGCCGCTAAATTTTTTTGATTTTAA 540
DB 481 CCAAGTAGCTGGGACTACAGGCGCCGCACTACGCGCCGCTAAATTTTTTTGATTTTAA 540
QY 541 GTAGAGACGGGGTTTACCGCTTTTACCGGGAATGCGCTGATCTCTGACCTGTGATCC 600
DB 541 GTAGAGACGGGGTTTACCGCTTTTACCGGGAATGCGCTGATCTCTGACCTGTGATCC 600
QY 601 GCCCGGCTGGCTCCCAAAGTCTGGGATTA CAGGCGTGAAGCCAGCGCCGCGCATG 660
DB 601 GCCCGGCTGGCTCCCAAAGTCTGGGATTA CAGGCGTGAAGCCAGCGCCGCGCATG 660
QY 661 ATCATCTTCTTGACTATGCTGATGTGA CAGTAACCTAAAGCCATCAGACTCTACCTTTA 720
DB 661 ATCATCTTCTTGACTATGCTGATGTGA CAGTAACCTAAAGCCATCAGACTCTACCTTTA 720
QY 721 AATATGCAAGTTGGGCGCAGGCAACCGGCTCAATGCTGTAAATCCAGCACTTTGGAGGC 780
DB 721 AATATGCAAGTTGGGCGCAGGCAACCGGCTCAATGCTGTAAATCCAGCACTTTGGAGGC 780
QY 781 AGAGGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCCAGCTGGGCAACATGTGAAG 840
DB 781 AGAGGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCCAGCTGGGCAACATGTGAAG 840
QY 841 CTCTGCTCTTAACTAAAAAATCAATCAACGCGGGTGTCTGTGGGCGAC 900
DB 841 CTCTGCTCTTAACTAAAAAATCAATCAACGCGGGTGTCTGTGGGCGAC 900
QY 901 AACTGTAATCCAGCTATGCTGAGGCTGAGCGAGAGTCACTTGAACCTTGGAGGCG 960
DB 901 AACTGTAATCCAGCTATGCTGAGGCTGAGCGAGAGTCACTTGAACCTTGGAGGCG 960
QY 961 GAGGTGCAAGTGGCGGAGATCAATCAACGCGCTCAGCTGGGCGAGAGAGAGCT 1020
DB 961 GAGGTGCAAGTGGCGGAGATCAATCAACGCGCTCAGCTGGGCGAGAGAGAGCT 1020
QY 1021 CTGCTCAAAATTAATTAATTAACAAACGAAAGCTTGTGTCTTAGTTATATCT 1080
DB 1021 CTGCTCAAAATTAATTAATTAACAAACGAAAGCTTGTGTCTTAGTTATATCT 1080
QY 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGTAAATTAAGAAAAATTAATGGGCGCA 1140
DB 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGTAAATTAAGAAAAATTAATGGGCGCA 1140
QY 1141 AGAATCTAAGGTATTTTGA CAAATCATTCAGAACCTTTAAAAAAGAAATCA CAGA 1200
DB 1141 AGAATCTAAGGTATTTTGA CAAATCATTCAGAACCTTTAAAAAAGAAATCA CAGA 1200
QY 1201 GGCATTAAGAGACAGGAGGAGACAGGAGACAGAAA CACTGTGTGCCCAAGAGAA CAAA 1260
DB 1201 GGCATTAAGAGACAGGAGGAGACAGGAGACAGAAA CACTGTGTGCCCAAGAGAA CAAA 1260
QY 1261 ACAAGGCTCTTAAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 ACAAGGCTCTTAAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

QY 1321 AAAAAACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGAGAGGATAGAAAG 1380
DB 1321 AAAAAACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGAGAGGATAGAAAG 1380
QY 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTTGAAGAGACTAGAAAGATTAGCCGAGGGA 1440
DB 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTTGAAGAGACTAGAAAGATTAGCCGAGGGA 1440
QY 1441 GAAACACAGAGAGATGAGAAAGAACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGAG 1500
DB 1441 GAAACACAGAGAGATGAGAAAGAACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGAG 1500
QY 1501 AGTATGAGGGTGAACAGACAGTGTGTGATAGCAAAATGACAGAGAAAGAAAGCAAGCA 1560
DB 1501 AGTATGAGGGTGAACAGACAGTGTGTGATAGCAAAATGACAGAGAAAGAAAGCAAGCA 1560
QY 1561 TCCAGGCGCCAAAGAAATAGTGAACCCAGAGTTGGTGAAGAACCAAGTCTTTAAGCTGGGG 1620
DB 1561 TCCAGGCGCCAAAGAAATAGTGAACCCAGAGTTGGTGAAGAACCAAGTCTTTAAGCTGGGG 1620
QY 1621 AGGCAAGGAGAGGGGCTGGCTGGGCTCCGAGAACCCCTCCCATTCCTCGGGCCAGGGAG 1680
DB 1621 AGGCAAGGAGAGGGGCTGGCTGGGCTCCGAGAACCCCTCCCATTCCTCGGGCCAGGGAG 1680
QY 1681 GTAGGAGATGACATTCGGAATGGGTGGGGGTGTCTTGGGGGTGAGATAGAGGAGAGCA 1740
DB 1681 GTAGGAGATGACATTCGGAATGGGTGGGGGTGTCTTGGGGGTGAGATAGAGGAGAGCA 1740
QY 1741 GGAAGAGCTATTGCTAAGGCCCGATAGGCACTCATTTGCCGGGAAATGTGCCCAAGGAG 1800
DB 1741 GGAAGAGCTATTGCTAAGGCCCGATAGGCACTCATTTGCCGGGAAATGTGCCCAAGGAG 1800
QY 1801 CAGTGGGTGTATTAATCAAGGCCCGGTGCCAGAGCCAGAGAGGAGGAGGAGGAGCAAG 1860
DB 1801 CAGTGGGTGTATTAATCAAGGCCCGGTGCCAGAGCCAGAGAGGAGGAGGAGGAGGAGCAAG 1860
QY 1861 AGGCAACAGGCTGAGAAAGTCTGCGCTGAGCTGAGGAGGAGAAATCCCAACCCCTACCTGG 1920
DB 1861 AGGCAACAGGCTGAGAAAGTCTGCGCTGAGCTGAGGAGGAGAAATCCCAACCCCTACCTGG 1920
QY 1921 GGGACAGGCGAAGTGAACCTGTGAGGGTGGCTCAGCAGGACAGGAGAGAGAGAGTGTCT 1980
DB 1921 GGGACAGGCGAAGTGAACCTGTGAGGGTGGCTCAGCAGGACAGGAGAGAGAGAGTGTCT 1980
QY 1981 GTGCGTCTGACCCCAACATCTTTCTGTGCCCTCTTGCCCTGTGTGAGGCTGTAGA 2040
DB 1981 GTGCGTCTGACCCCAACATCTTTCTGTGCCCTCTTGCCCTGTGTGAGGCTGTAGA 2040
QY 2041 CTCTATCTTCTGAATTTCTATAGTGCCTGAGTCTCAGCGCAGTGCAGATGTGCCCCCTG 2100
DB 2041 CTCTATCTTCTGAATTTCTATAGTGCCTGAGTCTCAGCGCAGTGCAGATGTGCCCCCTG 2100
QY 2101 CTGTGTGTTCTCTCTTCACTGTGGGAAATTAAGTGAAGGAGAGGAGAGAGTGGTTAAG 2160
DB 2101 CTGTGTGTTCTCTCTTCACTGTGGGAAATTAAGTGAAGGAGAGAGGAGAGAGTGGTTAAG 2160
QY 2161 GCTCCCGGATGAGCTGGGCTGCCCAACCTCTGACATTCCTCCATTCAGAGTCAAGCGGC 2220
DB 2161 GCTCCCGGATGAGCTGGGCTGCCCAACCTCTGACATTCCTCCATTCAGAGTCAAGCGGC 2220
QY 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTCTGTCTCTGATCA CAGCCTTGGCTT 2280
DB 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTCTGTCTCTGATCA CAGCCTTGGCTT 2280
QY 2281 CTGGGGGTCA CAGGTAA CAGAACTCTGTGGGTGGAGAGGTTTGGGATTTGGAGACCTGT 2340
DB 2281 CTGGGGGTCA CAGGTAA CAGAACTCTGTGGGTGGAGAGGTTTGGGATTTGGGAGACCTGT 2340
QY 2341 CTCTCGGACCTAAGAGCGCTGTCTCCCTGGGGGAACTGTGTGAGCGCTGGGCAATGACCTCGG 2400
DB 2341 CTCTCGGACCTAAGAGCGCTGTCTCCCTGGGGGAACTGTGTGAGCGCTGGGCAATGACCTCGG 2400
QY 2401 GACCGGCTGAATGTAGTCTCTGTCTGTACTTGTGTGTGTGAGTGTATGTGGCCCTGT 2460

| | | | |
|----|------|---|------|
| Db | 2401 | GACGGGGTGAATGTAAGTCTCTGTGTACTTGGGTGTGGATCGATGTGGCCCTGT | 2466 |
| Oy | 2461 | GACTGCCACGGGTGTGTCTGGGGAGAGGGAGATGCCCTTTCCCAATACAGGTGACTGTCCG | 2520 |
| Db | 2461 | GACTGCCACGGGTGTGTCTGGGGAGAGGGAGATGCCCTTTCCCAATACAGGTGACTGTCCG | 2520 |
| Oy | 2521 | CAGGTGGCACTGACCCCTTTAGAGCGTGTGTGTGGTTTGTGTGATTTGTGTGACTTTAAG | 2580 |
| Db | 2521 | CAGGTGGCACTGACCCCTTTAGAGCGTGTGTGTGGTTTGTGTGATTTGTGTGACTTTAAG | 2580 |
| Oy | 2581 | ATTGTGTGGCTCCACAGCTGTGTGGGTGAATGCAATGACACTGGGGGTTCACCTGT | 2640 |
| Db | 2581 | ATTGTGTGGCTCCACAGCTGTGTGGGTGAATGCAATGACACTGGGGGTTCACCTGT | 2640 |
| Oy | 2641 | GTGTGTGGCTGTGTGTGTGACTTTGGCAATTTGATATGTACTGACAGTATCTGCAATTCCTG | 2700 |
| Db | 2641 | GTGTGTGGCTGTGTGTGTGACTTTGGCAATTTGATATGTACTGACAGTATCTGCAATTCCTG | 2700 |
| Oy | 2701 | TCCCTGAGAGTCCCGGGATTGCGTGCACAAAAGGTGCATCAACAATGAAAAGCTGTGACT | 2760 |
| Db | 2701 | TCCCTGAGAGTCCCGGGATTGCGTGCACAAAAGGTGCATCAACAATGAAAAGCTGTGACT | 2760 |
| Oy | 2761 | GTGTGTGTCTGTGCAGGGCAATTATGTGATTTGTGTGTGCTGATGTGACGTTATGTGATGCCCTA | 2820 |
| Db | 2761 | GTGTGTGTCTGTGCAGGGCAATTATGTGATTTGTGTGTGCTGATGTGACGTTATGTGATGCCCTA | 2820 |
| Oy | 2821 | TTTGTGACCGGTGTGACTACCTGAGACTCTGTGTAGAGGGTGACTGTATGTGACTGTGTGTG | 2880 |
| Db | 2821 | TTTGTGACCGGTGTGACTACCTGAGACTCTGTGTAGAGGGTGACTGTGTATGTGACTGTGTGTG | 2880 |
| Oy | 2881 | TCTGTGTGAGGCGGTGTAAATGCTACTGTATGTGTGTGATGTGTGACGTGTGTGTCTTGAGT | 2940 |
| Db | 2881 | TCTGTGTGAGGCGGTGTAAATGCTACTGTATGTGTGTGATGTGTGACGTGTGTGTCTTGAGT | 2940 |
| Oy | 2941 | TTTGTGTCTGTGCTGTGAGAGGATAGAGGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC | 3000 |
| Db | 2941 | TTTGTGTCTGTGCTGTGAGAGGATAGAGGGGTGCAGGGGTAGCTATCTCTGGAGATGGGTGC | 3000 |
| Oy | 3001 | CAGGTGACTGACTTGTGACGTGTGTGCTGTGTGTGACAGAAAGATATGTGGCAGTCTGAACTC | 3060 |
| Db | 3001 | CAGGTGACTGACTTGTGACGTGTGTGCTGTGTGTGACAGAAAGATATGTGGCAGTCTGAACTC | 3060 |
| Oy | 3061 | TGTGTGACACAGGCATCTGTGTGTGTGACCTGAGACACTGTGTGATGAGGGTGTGGATGCC | 3120 |
| Db | 3061 | TGTGTGACACAGGCATCTGTGTGTGTGACCTGAGACACTGTGTGATGAGGGTGTGGATGCC | 3120 |
| Oy | 3121 | GCTAGGGCTGCCCGGGAGCGTGTGTACTGTGAGACAGACTGTATTTGTGCTGACCTGTGTG | 3180 |
| Db | 3121 | GCTAGGGCTGCCCGGGAGCGTGTGTACTGTGAGACAGACTGTATTTGTGCTGACCTGTGTG | 3180 |
| Oy | 3181 | GAGGCAACATGGGCGGTGTCTGTGAGAACAGCGGCGGTGTGTGGCTGTATCTGTGTGTGGC | 3240 |
| Db | 3181 | GAGGCAACATGGGCGGTGTCTGTGAGAACAGCGGCGGTGTGTGGCTGTATCTGTGTGTGGC | 3240 |
| Oy | 3241 | GCGTGTGTTCTTGTGGGTGTGATTCGTGAATGATGTGTGTGCTCAGAGGCATCAGCAAGGGTAA | 3300 |
| Db | 3241 | GCGTGTGTTCTTGTGGGTGTGATTCGTGAATGATGTGTGTGCTCAGAGGCATCAGCAAGGGTAA | 3300 |
| Oy | 3301 | GAAACGAGCCCGGGCCGGGTGGCTCAGCCTGTGAATCCAGCCCTTTTGGAGAGCCGAGGCA | 3360 |
| Db | 3301 | GAAACGAGCCCGGGCCGGGTGGCTCAGCCTGTGAATCCAGCCCTTTTGGAGAGCCGAGGCA | 3360 |
| Oy | 3361 | GGCGGATCACCTGAGTGTGGAGATCGAGGCCAGCTTACCAACATGGAGAAACCCCGCT | 3420 |
| Db | 3361 | GGCGGATCACCTGAGTGTGGAGATCGAGGCCAGCTTACCAACATGGAGAAACCCCGCT | 3420 |
| Oy | 3421 | CTACTTAAATAATCAAAATAATTAGCTGTGTGTGTGGTGCAGTCTGTATCCACGTACTC | 3480 |
| Db | 3421 | CTACTTAAATAATCAAAATAATTAGCTGTGTGTGTGGTGCAGTCTGTATCCACGTACTC | 3480 |
| Oy | 3481 | GGGAGACTGGGGCAGAAAAATCGTTGAACCCGGAGGTGTGAGGTTGCGTGTGACCGAGA | 3540 |

| | | | |
|----|------|---|------|
| Db | 3441 | GGAGAGCTGGGGCAGAAAAATCGCTTGAAACCCGGGAGGTGAGAGTTGCGGTGACCCAGA | 3540 |
| Qy | 3541 | TCGGCCATTGCTACCTCCAGCCTGGGCAACAAGCGAAATCTCGTCTTGAAAGAAAAA | 3600 |
| Db | 3541 | TCGGCCATTGCTACCTCCAGCCTGGGCAACAAGCGAAATCTCGTCTTGAAAGAAAAA | 3600 |
| Qy | 3601 | GAIAAAAAAAAAAGGTTAAGAACCACTGAATGGGCAACGGGAGACTGATGATGGAATTGGGGC | 3666 |
| Db | 3601 | GAIAAAAAAAAAAGGTTAAGAACCACTGAATGGGCAACGGGAGACTGATGATGGAATTGGGGC | 3666 |
| Qy | 3661 | ATGCATGAGTCTGTAAGGTCTGTGTGTGAGAGAGAGATTTGAACAGGATTGAAAGGCAT | 3720 |
| Db | 3661 | ATGCATGAGTCTGTAAGGTCTGTGTGTGAGAGAGAGATTTGAACAGGATTGAAAGGCAT | 3720 |
| Qy | 3721 | GTTTTTCATCTGAGAAATTCAGAAACCTAGAGCTGTCTTCCCTCCATGTGGCCCCCTAAG | 3780 |
| Db | 3721 | GTTTTTCATCTGAGAAATTCAGAAACCTAGAGCTGTCTTCCCTCCATGTGGCCCCCTAAG | 3780 |
| Qy | 3781 | CTGAGCCCTTCTTCTCTGTGCTCTGCTCTTGGAACCTTAGCTCCGCCCATGAGCTTGACC | 3840 |
| Db | 3781 | CTGAGCCCTTCTTCTCTGTGCTCTGCTCTTGGAACCTTAGCTCCGCCCATGAGCTTGACC | 3840 |
| Qy | 3841 | CCACTCCCTTCCCTCAACCAAGCCGCTAGGCAACTGTATGGAACCCCGCTAAGGCGCA | 3900 |
| Db | 3841 | CCACTCCCTTCCCTCAACCAAGCCGCTAGGCAACTGTATGGAACCCCGCTAAGGCGCA | 3900 |
| Qy | 3901 | CACCCCTTTGGGCGCAGGCTCCACCCCTTAATCTGTGGGTACTTCTAGAAACCCCTCTCA | 3960 |
| Db | 3901 | CACCCCTTTGGGCGCAGGCTCCACCCCTTAATCTGTGGGTACTTCTAGAAACCCCTCTCA | 3960 |
| Qy | 3961 | AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTGCCAGCTGGAG | 4020 |
| Db | 3961 | AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTGCCAGCTGGAG | 4020 |
| Qy | 4021 | TGCAGTGCGGATCTGGGCTCACTGGAACCTGCGCTCCAGGTTCAAGGATTTCTGT | 4080 |
| Db | 4021 | TGCAGTGCGGATCTGGGCTCACTGGAACCTGCGCTCCAGGTTCAAGGATTTCTGT | 4080 |
| Qy | 4081 | GCCTCAACCTCTGTAGTAGCTGGATTACAGGTGCGGCGCACACAGCTGGCTAATTTTT | 4140 |
| Db | 4081 | GCCTCAACCTCTGTAGTAGCTGGATTACAGGTGCGGCGCACACAGCTGGCTAATTTTT | 4140 |
| Qy | 4141 | GTGCTTTTAGTAGAGACAGGTTTCACTTGTGGCAAGCTGGTCTCAAACTCCAAAC | 4200 |
| Db | 4141 | GTGCTTTTAGTAGAGACAGGTTTCACTTGTGGCAAGCTGGTCTCAAACTCCAAAC | 4200 |
| Qy | 4201 | TCAAGTATTCGGCCCACTCGGCTCCCAAGTGTGAGGTTTAAAGCGTAGGCAACCGC | 4260 |
| Db | 4201 | TCAAGTATTCGGCCCACTCGGCTCCCAAGTGTGAGGTTTAAAGCGTAGGCAACCGC | 4260 |
| Qy | 4261 | CCCCAGCCCAAGTCAGAGCTCTTATAGGAGACTTAACATGTAAACCTGACCCCTGGCC | 4320 |
| Db | 4261 | CCCCAGCCCAAGTCAGAGCTCTTATAGGAGACTTAACATGTAAACCTGACCCCTGGCC | 4320 |
| Qy | 4321 | CTTAACTAAGTCAATTCAAAACCCCTTCTGCTCAGACCCCTGACCCCATCTGAGGCC | 4380 |
| Db | 4321 | CTTAACTAAGTCAATTCAAAACCCCTTCTGCTCAGACCCCTGACCCCATCTGAGGCC | 4380 |
| Qy | 4381 | TGACCCCCACTTCTTGAAGCAGTTCATCCCTTAAAGCCGTGTCTCCCTCCCACTCCGAG | 4440 |
| Db | 4381 | TGACCCCCACTTCTTGAAGCAGTTCATCCCTTAAAGCCGTGTCTCCCTCCCACTCCGAG | 4440 |
| Qy | 4441 | GCTCAGAGCCCAAGGCTTTGGCACTAACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT | 4500 |
| Db | 4441 | GCTCAGAGCCCAAGGCTTTGGCACTAACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAAT | 4500 |
| Qy | 4501 | TTTAACTCTCAATGTAGTCTTGAAGCAATTCAGGAATCTGTAGGTCCAGTTAGTCCA | 4560 |
| Db | 4501 | TTTAACTCTCAATGTAGTCTTGAAGCAATTCAGGAATCTGTAGGTCCAGTTAGTCCA | 4560 |
| Qy | 4561 | GTAACCTTACTGTAGGCTTGCTGTCTTGAAGCTTGAAGCTTGAGGCTTGAGAGGTGCCA | 4620 |
| Db | 4561 | GTAACCTTACTGTAGGCTTGCTGTCTTGAAGCTTGAAGCTTGAGGCTTGAGAGGTGCCA | 4620 |

QY 4621 CTCTTATTCGAGGCGCTGCGCCCTGAGCATGTCAGACACCCCTTAGCT 4680
DB 4621 CTCTTATTCGAGGCGCTGCGCCCTGAGCATGTCAGACACCCCTTAGCT 4680
QY 4681 GGTCTGCGCTTGTGAGTGTGAACCCACCCCGAGCCCGCTTGAAGCCCGCC 4740
DB 4681 GGTCTGCGCTTGTGAGTGTGAACCCACCCCGAGCCCGCTTGAAGCCCGCC 4740
QY 4741 CAACCATTTTCGCTTCCAGAGCATGTTCTGCGCAACATGATGTTCTGTGACACC 4800
DB 4741 CAACCATTTTCGCTTCCAGAGCATGTTCTGCGCAACATGATGTTCTGTGACACC 4800
QY 4801 CCTCTAACACCGTCCCTGTGAGAGCAACAGGAGCTGGAGCTGGGGCCGGGAGAGC 4860
DB 4801 CCTCTAACACCGTCCCTGTGAGAGCAACAGGAGCTGGAGCTGGGGCCGGGAGAGC 4860
QY 4861 CCGGTGTGAGTGAACAGAGAGCGCATCATCATGATTCGAGTCGATATGACACC 4920
DB 4861 CCGGTGTGAGTGAACAGAGAGCGCATCATCATGATTCGAGTCGATATGACACC 4920
QY 4921 AGCGGTGAGAGCGCGCTGTGTAAGGCCCAACAGCTTACTGCGGGCGGTGTGG 4980
DB 4921 AGCGGTGAGAGCGCGCTGTGTAAGGCCCAACAGCTTACTGCGGGCGGTGTGG 4980
QY 4981 TGCATTCACAGTGGCTGCTC 5000
DB 4981 TGCATTCACAGTGGCTGCTC 5000

RESULT 2
CQ874885 11570 bp DNA linear PAT 27-SBP-2004
LOCUS Sequence 3 from Patent WO2004077060.
DEFINITION CQ874885
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 Diamandis, E.P. and Petraki, C.
Assay for detection of renal cell carcinoma
Patent: WO 2004077060-A 3 10-SBP-2004;
Mount Sinai Hospital (CA)
Location/Qualifiers

FEATURES
source 1..11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCAGAGTGAAGCAAGAGAGTGAAGCTCCCTGCAAGAGGCTTGAATC 60
DB 1 GGGCCAGAGTGAAGCAAGAGAGTGAAGCTCCCTGCAAGAGGCTTGAATC 60
QY 61 TCCCTGCTTAAATGTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 61 TCCCTGCTTAAATGTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 AAGAAAG 180
DB 121 AAGAAAG 180
QY 181 ACAAGAGAGCTGGAGACAGAGGAGACAGAGAGTCAAGAGAGAGAGAGAGAG 240
DB 181 ACAAGAGAGCTGGAGACAGAGGAGACAGAGAGTCAAGAGAGAGAGAGAGAG 240

QY 241 ACACAATGAGACACAGAGGTGAAGAAAGAGAGATTAAAGAGTCCAGATACACG 300
DB 241 ACACAATGAGACACAGAGGTGAAGAAAGAGAGATTAAAGAGTCCAGATACACG 300
QY 301 AAAAGGAGAGAGACAGATTTCAGGGTGTGTATGATCATCTTCTTTTTTTTTT 360
DB 301 AAAAGGAGAGAGACAGATTTCAGGGTGTGTATGATCATCTTCTTTTTTTTTT 360
QY 361 TTTTTTTTTTTTTTGAAGCGAGGTGTGCTGTGTGCGCCAGGCTGAGATGAGGGG 420
DB 361 TTTTTTTTTTTTTTGAAGCGAGGTGTGCTGTGTGCGCCAGGCTGAGATGAGGGG 420
QY 421 GATCTCGCTCAGTCGAAAGCTCCGCTCCCGGGTTTCAAGCAATCTCTGCTCAGCTC 480
DB 421 GATCTCGCTCAGTCGAAAGCTCCGCTCCCGGGTTTCAAGCAATCTCTGCTCAGCTC 480
QY 481 CCAAGTACTGGAGCTACAGGCGCCCGCCTACAGCCCGGCTAAATTTTTTGTATTTTA 540
DB 481 CCAAGTACTGGAGCTACAGGCGCCCGCCTACAGCCCGGCTAAATTTTTTGTATTTTA 540
QY 541 GTAGAGAGCGGGGTTTCAACGTTTGAAGCGGAGTGGCTGATCTGCTGACCTGATATC 600
DB 541 GTAGAGAGCGGGGTTTCAACGTTTGAAGCGGAGTGGCTGATCTGCTGACCTGATATC 600
QY 601 GCCCGCTTCGCGCTCCCAAAGTCTGAGATTACAGGCGTGAACACCGCGCCGAGC 660
DB 601 GCCCGCTTCGCGCTCCCAAAGTCTGAGATTACAGGCGTGAACACCGCGCCGAGC 660
QY 661 ATCATCTTCTTGAATGATGTGATGTGACAGATGTAAGCCATGAGCTTACCTTTA 720
DB 661 ATCATCTTCTTGAATGATGTGATGTGACAGATGTAAGCCATGAGCTTACCTTTA 720
QY 721 AATATGAGTTTGGGCGAGGACCGTGGCTACGCTGTAATTCAGACCTTGGAGGCG 780
DB 721 AATATGAGTTTGGGCGAGGACCGTGGCTACGCTGTAATTCAGACCTTGGAGGCG 780
QY 781 AGAGGTGGGTGAATCACTTGAAGCCAGAGATTGAGACCAAGCTGCGCAATGTGAAA 840
DB 781 AGAGGTGGGTGAATCACTTGAAGCCAGAGATTGAGACCAAGCTGCGCAATGTGAAA 840
QY 841 CTCTGCTTTTACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
DB 841 CTCTGCTTTTACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
QY 901 ACCGTGAATCCAGCATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 901 ACCGTGAATCCAGCATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 GAGGTTGAGTGGGCGAGATCATCATCCGCTCCAGCTGAGCGACAGAGCAAGACT 1020
DB 961 GAGGTTGAGTGGGCGAGATCATCATCCGCTCCAGCTGAGCGACAGAGCAAGACT 1020
QY 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
DB 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
QY 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1141 AGAAGCTTAAGGTATATTGCAATCATTCAGAACTTTTAAAAAAGAAATCAGAG 1200
DB 1141 AGAAGCTTAAGGTATATTGCAATCATTCAGAACTTTTAAAAAAGAAATCAGAG 1200
QY 1201 GGCATTAAGAAAGCAGAGGAGGAGCAGGAGAGCAGAAACCTGTGGCCCAAGGAGAA 1260
DB 1201 GGCATTAAGAAAGCAGAGGAGGAGCAGGAGAGCAGAAACCTGTGGCCCAAGGAGAA 1260
QY 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AAAAAAG 1380

1321 AAAAAGCAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGATATAGAAAGA 1380
1381 GAGAGAGGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCGAAAAGATAGCCGAGGA 1440
1381 GAGAGAGGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCGAAAAGATAGCCGAGGA 1440
1441 GAACCCAGAGAGATGGAGAGAGACTGTGAAAAAACAGAGACAAAGATGGAAAGAG 1500
1441 GAACCCAGAGAGATGGAGAGAGACTGTGAAAAAACAGAGACAAAGATGGAAAGAG 1500
1501 AGTATGAGGGGTGACAGACAGATGTGGAAATGAGCAAAATGACAGAGAAAGCAAGCA 1560
1501 AGTATGAGGGGTGACAGACAGATGTGGAAATGAGCAAAATGAGCAAGAAAGCAAGCA 1560
1561 TCACGCGCCCAAGAAATAGTACCCAGATTGGTGAAGCCAGATCTTTAAGCTGGGG 1620
1561 TCACGCGCCCAAGAAATAGTACCCAGATTGGTGAAGCCAGATCTTTAAGCTGGGG 1620
1621 AGGACAGGAAAGGGGCTGGCTTCCGAGACCCCTCCCAATTCTCCGGGCGAGGGAG 1680
1621 AGGACAGGAAAGGGGCTGGCTTCCGAGACCCCTCCCAATTCTCCGGGCGAGGGAG 1680
1681 GTAGGAGTGAATTCCGAGCTGGTGGGGGTGCTCTGGGGGTGAGATAGGGGGAGCA 1740
1681 GTAGGAGTGAATTCCGAGCTGGTGGGGGTGCTCTGGGGGTGAGATAGGGGGAGCA 1740
1741 GAGAGAGCTATTGCTTAAGGCCGAGATAGGACCTCATTTGCCGGAAATGTGCCAGGAG 1800
1741 GAGAGAGCTATTGCTTAAGGCCGAGATAGGACCTCATTTGCCGGAAATGTGCCAGGAG 1800
1801 CAGTGGTGGTTTAATCTAGAGCCGGTGGCCAGAGCCAGAGAGAGAGAGAGAGAGCA 1860
1801 CAGTGGTGGTTTAATCTAGAGCCGGTGGCCAGAGCCAGAGAGAGAGAGAGAGAGCA 1860
1801 CAGTGGTGGTTTAATCTAGAGCCGGTGGCCAGAGCCAGAGAGAGAGAGAGAGAGCA 1860
1861 AGGACAGGCGCTGAGAGAGTCTGGGCTGAGCTGAGCAAAATCCCCACCCCTACTGCG 1920
1861 AGGACAGGCGCTGAGAGAGTCTGGGCTGAGCTGAGCAAAATCCCCACCCCTACTGCG 1920
1861 AGGACAGGCGCTGAGAGAGTCTGGGCTGAGCTGAGCAAAATCCCCACCCCTACTGCG 1920
1921 GGGACAGGCGAGATGAGAGCTGGTGAAGGTGGCTCAGAGAGAGAGAGAGAGGTGTCT 1980
1921 GGGACAGGCGAGATGAGAGCTGGTGAAGGTGGCTCAGAGAGAGAGAGAGAGGTGTCT 1980
1981 GTCGCTCTGCAACCATCTTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 2040
1981 GTCGCTCTGCAACCATCTTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 2040
2041 CTCCTATTCTTAATTTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2041 CTCCTATTCTTAATTTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2101 CTTTGTGTTCTCTCTAATCTGAGGAAATTAAGTAAAGGAGAGAGAGAGAGGTGTTAAG 2160
2101 CTTTGTGTTCTCTCTAATCTGAGGAAATTAAGTAAAGGAGAGAGAGAGAGGTGTTAAG 2160
2161 GCTCCCCGAGATGCGCTGGGCTGCCAACCCTGAGCAATTCCTCAATCCAGGTGAGGGCC 2220
2161 GCTCCCCGAGATGCGCTGGGCTGCCAACCCTGAGCAATTCCTCAATCCAGGTGAGGGCC 2220
2221 ATGGCTACAG 2280
2221 ATGGCTACAG 2280
2281 CTGGGGGTGACAGGTAAACAGAACTCTGGGGGTGAGAGAGAGAGAGAGAGAGAGAG 2340
2281 CTGGGGGTGACAGGTAAACAGAACTCTGGGGGTGAGAGAGAGAGAGAGAGAGAGAG 2340
2341 CTCTGGGGGACAG 2400
2341 CTCTGGGGGACAG 2400
2401 GACCGGGTGAATGAGTCTCTGTCTGTATCTTGTGTTGCGATGCTATGTGGCCCTGT 2460
2401 GACCGGGTGAATGAGTCTCTGTCTGTATCTTGTGTTGCGATGCTATGTGGCCCTGT 2460

2401 GACCGGGTGAATGAGTCTCTGTCTGTATCTTGTGTTGCGATGCTATGTGGCCCTGT 2460
2461 GACTGCCAGGTGTGTGTGGGAGGGGAGATGCTTTTCCATATGAGAGAGAGAGAGAG 2520
2461 GACTGCCAGGTGTGTGTGGGAGGGGAGATGCTTTTCCATATGAGAGAGAGAGAGAG 2520
2521 CAGGTGGACAG 2580
2521 CAGGTGGACAG 2580
2581 ATTGTGTGTGCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2581 ATTGTGTGTGCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2641 GT 2700
2641 GT 2700
2701 TCCCTGAGGTTCCGGGATTTGGCGTCAACAAAGTGTGATCAACATGAGAAAGCTGTGACT 2760
2701 TCCCTGAGGTTCCGGGATTTGGCGTCAACAAAGTGTGATCAACATGAGAAAGCTGTGACT 2760
2761 GT 2820
2761 GT 2820
2821 TTTGTGACCGT 2880
2821 TTTGTGACCGT 2880
2881 TCTGTGTGAGCCGTTAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
2881 TCTGTGTGAGCCGTTAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
2941 TTTGTGTCTTCTGCTGTGAGGAGATAGAGGTGTGACAGGGGTGACTATCTTGGAGAGATGGGTGC 3000
2941 TTTGTGTCTTCTGCTGTGAGGAGATAGAGGTGTGACAGGGGTGACTATCTTGGAGAGATGGGTGC 3000
2941 TTTGTGTCTTCTGCTGTGAGGAGATAGAGGTGTGACAGGGGTGACTATCTTGGAGAGATGGGTGC 3000
3001 CAGGTGACAG 3060
3001 CAGGTGACAG 3060
3061 TGTGACACAG 3120
3061 TGTGACACAG 3120
3121 GCTAGGCTGCCGGAGCGT 3180
3121 GCTAGGCTGCCGGAGCGT 3180
3181 GAGGCAACATGGGCGT 3240
3181 GAGGCAACATGGGCGT 3240
3241 GGT 3300
3241 GGT 3300
3301 GAACCAAGCGGGCGGAGT 3360
3301 GAACCAAGCGGGCGGAGT 3360
3301 GAACCAAGCGGGCGGAGT 3360
3361 GGGGATTCACCTGAGGTGCGGAGATCGAGAGCAGAGCTGACCAAGAGAGAAACCCGCTGT 3420
3361 GGGGATTCACCTGAGGTGCGGAGATCGAGAGCAGAGCTGACCAAGAGAGAAACCCGCTGT 3420
3421 CTAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 3480
3421 CTAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 3480
3481 GGGAGACTGGGCGAGAAATATGCTTTGAACCCGGGAGGTGTGTGTGTGTGTGTGTGTGT 3540
3481 GGGAGACTGGGCGAGAAATATGCTTTGAACCCGGGAGGTGTGTGTGTGTGTGTGTGTGT 3540

3541 TCGCGCATTTGACATCCAGCTGGGCAAGAGCGAAACTCCGTCTCGAAAAGAAAAA 3600
3541 TCGCGCATTTGACATCCAGCTGGGCAAGAGCGAAACTCCGTCTCGAAAAGAAAAA 3600
3601 GAAAAAAGAGGTAAGAACAGTGAATGGGCAAGAGAGACTGATGATGATGGGCG 3660
3601 GAAAAAAGAGGTAAGAACAGTGAATGGGCAAGAGAGACTGATGATGATGGGCG 3660
3661 ATGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
3661 ATGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
3721 GTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
3721 GTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
3781 CTGAGCGCTTTCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3840
3781 CTGAGCGCTTTCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3840
3841 CCACCTCTTCTCTCAACAGCGCGCTAGGCGAGACTAGTGAACCGCGCTAAGGCCA 3900
3841 CCACCTCTTCTCTCAACAGCGCGCTAGGCGAGACTAGTGAACCGCGCTAAGGCCA 3900
3901 CACCCCTTTGGGCAAGGCTCAACCCCTATTCGTGGGATCCTTCAAGAACCCCTTCA 3960
3901 CACCCCTTTGGGCAAGGCTCAACCCCTATTCGTGGGATCCTTCAAGAACCCCTTCA 3960
3961 AGTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4020
3961 AGTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4020
4021 TGCAAGTGGGATGATCTCGGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCT 4080
4021 TGCAAGTGGGATGATCTCGGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCT 4080
4081 GCTCCACCTCTGAGTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4140
4081 GCTCCACCTCTGAGTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4140
4081 GCTCCACCTCTGAGTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4140
4141 GTGTCTTTAGTGAAGCAGGAGTTCACCTTTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4200
4141 GTGTCTTTAGTGAAGCAGGAGTTCACCTTTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4200
4141 GTGTCTTTAGTGAAGCAGGAGTTCACCTTTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4200
4201 TAGAGTATCCGCGCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAAGCAGCGC 4260
4201 TAGAGTATCCGCGCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAAGCAGCGC 4260
4261 CCCGAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4320
4261 CCCGAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4320
4261 CCCGAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4320
4321 CTAACTAACTCAATTCAAAACCTTCTGCTCCAGGCTGAGGCTGAGGCTGAGGCTGAG 4380
4321 CTAACTAACTCAATTCAAAACCTTCTGCTCCAGGCTGAGGCTGAGGCTGAGGCTGAG 4380
4321 CTAACTAACTCAATTCAAAACCTTCTGCTCCAGGCTGAGGCTGAGGCTGAGGCTGAG 4380
4381 TGAACCCACTTTCTTGAAGCAAGTTCATCCCTTAAGCCTCTGCTCTCCATCCCGAG 4440
4381 TGAACCCACTTTCTTGAAGCAAGTTCATCCCTTAAGCCTCTGCTCTCCATCCCGAG 4440
4441 GCTCCAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4500
4441 GCTCCAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4500
4441 GCTCCAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAACTCTGAACCTTGCC 4500
4501 TTTTACCTCACTGATGATCTGAGCAATTCAGGAATCTGTAGAGCTCAGTTAGAGTCA 4560
4501 TTTTACCTCACTGATGATCTGAGCAATTCAGGAATCTGTAGAGCTCAGTTAGAGTCA 4560
4501 TTTTACCTCACTGATGATCTGAGCAATTCAGGAATCTGTAGAGCTCAGTTAGAGTCA 4560
4561 GTTAACTCACTGAGCTGGGCTGTCTGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4620
4561 GTTAACTCACTGAGCTGGGCTGTCTGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4620

4621 CTCTATTCTCCAGGCGCTGCGCCCTGCGCCCTCAGCATGTGACACACCCCTTAAGT 4680
4621 CTCTATTCTCCAGGCGCTGCGCCCTGCGCCCTCAGCATGTGACACACCCCTTAAGT 4680
4681 GGTCTGGCTCTTTGATCTGAAACCCACCCAGCCCAAGCCCGCTTGAAGCCCGCC 4740
4681 GGTCTGGCTCTTTGATCTGAAACCCACCCAGCCCAAGCCCGCTTGAAGCCCGCC 4740
4741 CAACCAATTTCTGCTCCAGAGCATGTTCTGCGCAACATATGATGTTCTGTAACACC 4800
4741 CAACCAATTTCTGCTCCAGAGCATGTTCTGCGCAACATATGATGTTCTGTAACACC 4800
4801 CCTTAACACCGTCTCTTCTGAGCAACAGAGCTGAGGCTGAGGCTGAGGCTGAGGCT 4860
4801 CCTTAACACCGTCTCTTCTGAGCAACAGAGCTGAGGCTGAGGCTGAGGCTGAGGCT 4860
4861 CCGGTGGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4920
4861 CCGGTGGATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4920
4921 AGCGTGGAGGCGCGCTGTTGCTAAGGCCAACAGCTTACTGCGGCGGCTGTTGG 4980
4921 AGCGTGGAGGCGCGCTGTTGCTAAGGCCAACAGCTTACTGCGGCGGCTGTTGG 4980
4981 TGCAATCAACAGTGGCTGCTC 5000
4981 TGCAATCAACAGTGGCTGCTC 5000

RESULT 3
CO874961 11570 bp DNA linear PAT 27-SEP-2004
LOCUS CO874961
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION CO874961
VERSION CO874961.1 GI:52748060
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
source
1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGCCAGAGTGAAGCAAGAGAGTGAAGCTCCCTGCAAAAGTGGCTTGAATC 60
1 GGGCCAGAGTGAAGCAAGAGAGTGAAGCTCCCTGCAAAAGTGGCTTGAATC 60
61 TCCCTGCTTAATATGACAGGAGGAGGAGCAAGAAACAGGAAAGAGAGGAGGAGG 120
61 TCCCTGCTTAATATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
121 AAGAAAG 180
121 AAGAAAG 180
181 ACAGAGAGCTGGGAGCAGAGGAGCAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 240
181 ACAGAGAGCTGGGAGCAGAGGAGCAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 240
241 ACACAAATGAGAGCAGAGAGGTGTAAAGAAAGAGAGATTAACAGAGTCCAGATACAGC 300

Db 241 ACACAAATGGAGACAGAGGTGTAAAGAAAGAGATTAAACAGATCCAGATCAAGC 300
Qy 301 AAAGGGGAGAGACAGATTTCAGGGTGTCTATGATCATCTCTTTTTTTTTTTT 360
Db 301 AAAGGGGAGAGACAGATTTCAGGGTGTCTATGATCATCTCTTTTTTTTTTTT 360
Qy 361 TTTTTTTTTTTTTTTTGAAGAGAGTCTGTGTGCGCCAGGCTGAGAGTGGCGG 420
Db 361 TTTTTTTTTTTTTTTTGAAGAGAGTCTGTGTGCGCCAGGCTGAGAGTGGCGG 420
Qy 421 GATCTGGGCTCACTGCAAGCTCCGGCTCCGGGTTCAGGCATCTCTCTCAGCTC 480
Db 421 GATCTGGGCTCACTGCAAGCTCCGGCTCCGGGTTCAGGCATCTCTCTCAGCTC 480
Qy 481 CCAAGTAGCTGGAGCTACAGAGCGCCGCACTAGCGCCGCTAATTTTTTTTGA 540
Db 481 CCAAGTAGCTGGAGCTACAGAGCGCCGCACTAGCGCCGCTAATTTTTTTTGA 540
Qy 541 GTAGAGACGGGGTTTTCAACGTTTTTAAAGCGGAGTGGCTCGATCTCTGACCTG 600
Db 541 GTAGAGACGGGGTTTTCAACGTTTTTAAAGCGGAGTGGCTCGATCTCTGACCTG 600
Qy 601 GCCCGCTCGGCTCCCAAGTGTGGGATTACAGGCGTGAGCCACCGCCGCGCATG 660
Db 601 GCCCGCTCGGCTCCCAAGTGTGGGATTACAGGCGTGAGCCACCGCCGCGCATG 660
Qy 661 ATATCTTCTTGACTATGCTGATGTGAACAATCAATCAATCAATCAATCAAT 720
Db 661 ATATCTTCTTGACTATGCTGATGTGAACAATCAATCAATCAATCAATCAAT 720
Qy 721 AAATATGAGTTTGGGCGCAGGACCGTGGCTCATGCTCTAATTCAGACCTTTGG 780
Db 721 AAATATGAGTTTGGGCGCAGGACCGTGGCTCATGCTCTAATTCAGACCTTTGG 780
Qy 781 AGAGTGGGAGTAATCACTTGAAGGCAAGAGTTTGAAGCAAGCTTGGCAATG 840
Db 781 AGAGTGGGAGTAATCACTTGAAGGCAAGAGTTTGAAGCAAGCTTGGCAATG 840
Qy 841 CTCTGCTTTTAACTAAAAAATCAAGCGGGTGTCTGGGGCAC 900
Db 841 CTCTGCTTTTAACTAAAAAATCAAGCGGGTGTCTGGGGCAC 900
Qy 901 ACCGTGTAATCCAGCTATGCTGAGGCTGAGGCAAGAGTCACTTGAACCTTGA 960
Db 901 ACCGTGTAATCCAGCTATGCTGAGGCTGAGGCAAGAGTCACTTGAACCTTGA 960
Qy 961 GAGGTGCAAGTGGGCGAGATCATACCGCCCTCAGGCTGGGCGAGAGCAAGAT 1020
Db 961 GAGGTGCAAGTGGGCGAGATCATACCGCCCTCAGGCTGGGCGAGAGCAAGAT 1020
Qy 1021 CTGTCTAAATAAATAAATAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080
Db 1021 CTGTCTAAATAAATAAATAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1080
Qy 1081 AAAAAAATGCTGTCTAACAATAGCAAGAGTGAATTAAGAAAAAATTAATG 1140
Db 1081 AAAAAAATGCTGTCTAACAATAGCAAGAGTGAATTAAGAAAAAATTAATG 1140
Qy 1141 AGAAGCTAAGGTATATTTGACAAATCATTCAGAACTTTAAAAAAGAAAGAT 1200
Db 1141 AGAAGCTAAGGTATATTTGACAAATCATTCAGAACTTTAAAAAAGAAAGAT 1200
Qy 1201 GGCATAGAAAAGACAGGAGGAAAGAGAGACAGAAACACTGTGGCCCAAGAG 1260
Db 1201 GGCATAGAAAAGACAGGAGGAAAGAGAGACAGAAACACTGTGGCCCAAGAG 1260
Qy 1261 ACAAGGCTCTTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380

Db 1381 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 GAGAGAGGGGTGAGAGAGACACAGAGATTTTGAAGAGACTCAGAAAGATAG 1440
Db 1381 GAGAGAGGGGTGAGAGAGACACAGAGATTTTGAAGAGACTCAGAAAGATAG 1440
Qy 1441 GAAACACAGAGAGATGAGAAAGAGACTCTGAGAAAAAACAGAGACAAAGT 1500
Db 1441 GAAACACAGAGAGATGAGAAAGAGACTCTGAGAAAAAACAGAGACAAAGT 1500
Qy 1501 AGTATGAGAGGTGAACAGACAGTGTGTAATGACAAATTCAGAGAAAGCAAG 1560
Db 1501 AGTATGAGAGGTGAACAGACAGTGTGTAATGACAAATTCAGAGAAAGCAAG 1560
Qy 1561 TCCAGGCGCCAGAAATAGTGAACCCAGATTTGGTGAAGAACCATCTTAA 1620
Db 1561 TCCAGGCGCCAGAAATAGTGAACCCAGATTTGGTGAAGAACCATCTTAA 1620
Qy 1621 AGGCAAGGAAAGGGGCTGGCTGGCTCCGAGACCCCTCCCATTTCTCGGG 1680
Db 1621 AGGCAAGGAAAGGGGCTGGCTGGCTCCGAGACCCCTCCCATTTCTCGGG 1680
Qy 1681 GTAGGAGTGAACATTCGGAATGGAGCTGGAGGAGTCTCTGGGGTGAAGAT 1740
Db 1681 GTAGGAGTGAACATTCGGAATGGAGCTGGAGGAGTCTCTGGGGTGAAGAT 1740
Qy 1741 GGAAGAGCTATTGCTTAAGCCGAGATAGGACCTCATTTGCCCGGGAATG 1800
Db 1741 GGAAGAGCTATTGCTTAAGCCGAGATAGGACCTCATTTGCCCGGGAATG 1800
Qy 1801 CAGTGGGTGTTAATCTAGAGCCGGTGGTCCAGAGCCAGAGAGAGAGAGAG 1860
Db 1801 CAGTGGGTGTTAATCTAGAGCCGGTGGTCCAGAGCCAGAGAGAGAGAGAG 1860
Qy 1861 AGGCAAGGCTGAGAGAGTCTGGGCTGAGCTGGAGCAAAATCCCACTTAC 1920
Db 1861 AGGCAAGGCTGAGAGAGTCTGGGCTGAGCTGGAGCAAAATCCCACTTAC 1920
Qy 1921 GGAACAGGCGAGTGAAGCTGTGGTGAAGGTGTCTCAGCAGGCAAGGAGAG 1980
Db 1921 GGAACAGGCGAGTGAAGCTGTGGTGAAGGTGTCTCAGCAGGCAAGGAGAG 1980
Qy 1981 GTGCGTCTGACCCCATCTTTCTGTGCTCCCTCTTGGCCCTGTGAGAGCT 2040
Db 1981 GTGCGTCTGACCCCATCTTTCTGTGCTCCCTCTTGGCCCTGTGAGAGCT 2040
Qy 2041 CTCTATCTTCTGAATTTCTAATAGTGCCTGGTCTCAGCGCAGTGCAGTGC 2100
Db 2041 CTCTATCTTCTGAATTTCTAATAGTGCCTGGTCTCAGCGCAGTGCAGTGC 2100
Qy 2101 CTGTGGTTCCTCTTACCTGGGAAATTAAGTGAAGGAGAGAGAGAGAGAG 2160
Db 2101 CTGTGGTTCCTCTTACCTGGGAAATTAAGTGAAGGAGAGAGAGAGAGAG 2160
Qy 2161 GCTCCCGGAGTGGCTGGGCTGCCAACCCTGACATTCCTCCATCAGGTG 2220
Db 2161 GCTCCCGGAGTGGCTGGGCTGCCAACCCTGACATTCCTCCATCAGGTG 2220
Qy 2221 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGTCTGTATCAAGCT 2280
Db 2221 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGTCTGTATCAAGCT 2280
Qy 2281 CTGGGGGTCAAGGTAAACAGAACTTGGGGTGGAGGTTGTGGATTGGAGAG 2340
Db 2281 CTGGGGGTCAAGGTAAACAGAACTTGGGGTGGAGGTTGTGGATTGGAGAG 2340
Qy 2341 CTGTGGGCACTAGAGAGGCTGTCCCTGGGGAATGTGTGAGCTTGGGATGA 2400
Db 2341 CTGTGGGCACTAGAGAGGCTGTCCCTGGGGAATGTGTGAGCTTGGGATGA 2400
Qy 2401 GACCGGGTGAATGTAGTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 GACCGGGTGAATGTAGTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGT 2460

QY 2461 GACTGCCAGGTTGTGTGCGGGAGGGGATGCTTTTCCATATCAGGTGACCTGCGG 2520
DB 2461 GACTGCCAGGTTGTGTGCGGGAGGGGATGCTTTTCCATATCAGGTGACCTGCGG 2520
QY 2521 CAGGTGGACCTGACCTTTGAGGCTGTGTGTGTTTGTGTGATGTGTGTGATTTAA 2580
DB 2521 CAGGTGGACCTGACCTTTGAGGCTGTGTGTGTTTGTGTGATGTGTGTGATTTAA 2580
QY 2581 ATTGT 2640
DB 2581 ATTGT 2640
QY 2641 GT 2700
DB 2641 GT 2700
QY 2701 TCCCTGAGGTCCTGGGATGCGTGAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
DB 2701 TCCCTGAGGTCCTGGGATGCGTGAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
QY 2761 GT 2820
DB 2761 GT 2820
QY 2821 TTTGTGACCGT 2880
DB 2821 TTTGTGACCGT 2880
QY 2881 TCTGT 2940
DB 2881 TCTGT 2940
QY 2941 TTTGT 3000
DB 2941 TTTGT 3000
QY 3001 CAGGTGACTGACTTGT 3060
DB 3001 CAGGTGACTGACTTGT 3060
QY 3061 TGTGTGACACACGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
DB 3061 TGTGTGACACACGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
QY 3121 GCTAGGCTGCTGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
DB 3121 GCTAGGCTGCTGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
QY 3181 GAGGCAACATGAGCGGT 3240
DB 3181 GAGGCAACATGAGCGGT 3240
QY 3241 GCGT 3300
DB 3241 GCGT 3300
QY 3301 GAAACGAGCGCGGCGGT 3360
DB 3301 GAAACGAGCGCGGCGGT 3360
QY 3361 GCGGATGACCTGT 3420
DB 3361 GCGGATGACCTGT 3420
QY 3421 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3480
DB 3421 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3480
QY 3481 GGGAGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540
DB 3481 GGGAGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540

QY 3541 TCGCGCATTTGACCTTCAACCTGAGGCAACAGAGGAGGAGGAGGAGGAGGAGGAG 3600
DB 3541 TCGCGCATTTGACCTTCAACCTGAGGCAACAGAGGAGGAGGAGGAGGAGGAGGAG 3600
QY 3601 GAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
DB 3601 GAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
QY 3661 ATGCAATGATGT 3720
DB 3661 ATGCAATGATGT 3720
QY 3721 GTTTTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
DB 3721 GTTTTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
QY 3781 GTGAGCGCTTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3840
DB 3781 GTGAGCGCTTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3840
QY 3841 CCAACCTTCTTCTGCTGT 3900
DB 3841 CCAACCTTCTTCTGCTGT 3900
QY 3901 CACCCCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960
DB 3901 CACCCCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960
QY 3961 AGTCAGAGCTTTTGT 4020
DB 3961 AGTCAGAGCTTTTGT 4020
QY 4021 TGCAGT 4080
DB 4021 TGCAGT 4080
QY 4081 GCTTCACCTTCTGAGT 4140
DB 4081 GCTTCACCTTCTGAGT 4140
QY 4141 GTGTCTTTTGT 4200
DB 4141 GTGTCTTTTGT 4200
QY 4201 TGAAGT 4260
DB 4201 TGAAGT 4260
QY 4261 CCCAGCCCAAGT 4320
DB 4261 CCCAGCCCAAGT 4320
QY 4321 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4380
DB 4321 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4380
QY 4381 TGAACCACTTGT 4440
DB 4381 TGAACCACTTGT 4440
QY 4441 GCTTCAGCCCAAGT 4500
DB 4441 GCTTCAGCCCAAGT 4500
QY 4501 TTTTACCTTCAATGT 4560
DB 4501 TTTTACCTTCAATGT 4560
QY 4561 GTTAACCTTCAATGT 4620
DB 4561 GTTAACCTTCAATGT 4620
QY 4621 CTCTTATTTCTGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4680

| | | | |
|----|------|---|------|
| Db | 4621 | CTTATATCTCCAGGCGCTGCGCCCTGCGCCCTCCAGCATGTGACACCCACCCTTAACT | 4680 |
| Qy | 4681 | GGTCTGAGCCTCTTGTAGTCTGAAGAACCCAGCCAGCCAGCCGCGCTGTAGCCCGGCG | 4740 |
| Db | 4661 | GGTCTGAGCCTCTTGTAGTCTGAAGAACCCAGCCAGCCAGCCGCGCTGTAGCCCGGCG | 4740 |
| Qy | 4741 | CAACCATATTTCCGTTCCAGAGCATGTTCTGSCCAATATGATGTTTCTGTGACAC | 4800 |
| Db | 4741 | CAACCATATTTCCGTTCCAGAGCATGTTCTGSCCAATATGATGTTTCTGTGACAC | 4800 |
| Qy | 4801 | CCTCTAACACCGTGCCCTCTGCGAGCAACAGACCTGCGAGCTGCGGCCGCGGGAAGCG | 4860 |
| Db | 4801 | CCTCTAACACCGTGCCCTCTGCGAGCAACAGACCTGCGAGCTGCGGCCGCGGGAAGCG | 4860 |
| Qy | 4861 | CCCGGTGCGATGACAGCAGACAGCCGCATCATCAATGATTCGACTGCGATATGCACACC | 4920 |
| Db | 4861 | CCCGGTGCGATGACAGCAGACAGCCGCATCATCAATGATTCGACTGCGATATGCACACC | 4920 |
| Qy | 4921 | AGCGTGGCAGGCGCGCGCTGTTGCTTAAGGCCCAACAGCTCTACTGCGCGGCGGTTTGG | 4980 |
| Db | 4921 | AGCGTGGCAGGCGCGCGCTGTTGCTTAAGGCCCAACAGCTCTACTGCGCGGCGGTTTGG | 4980 |
| Qy | 4981 | TGdATCCAAGTGGCTGCTC 5000 | |
| Db | 4981 | TGdATCCAAGTGGCTGCTC 5000 | |

| | | | | | |
|------------|--|------------|-----|--------|-----------------|
| LOCUS | AF135028 | 11570 bp | DNA | linear | PRI 27-JUN-2000 |
| DEFINITION | Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds. | | | | |
| ACCESSION | AF135028 | | | | |
| VERSION | AF135028.1 | GI:4589282 | | | |
| KEYWORDS | . | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 11570) | | | | |
| TITLE | Yousef,G.M. and Diamandis,E.P. | | | | |
| JOURNAL | The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation | | | | |
| MEDLINE | J Biol. Chem. 274 (53), 37511-37516 (1999) | | | | |
| PUBMED | 20076408 | | | | |
| REFERENCE | 2 (bases 1 to 11570) | | | | |
| AUTHORS | Diamandis,E.P., Yousef,G.M., Luo,L.Y., Maglara,A. and Obiezu,C.V. | | | | |
| TITLE | The new human kallikrein gene family: implications in carcinogenesis | | | | |
| JOURNAL | Trends Endocrinol. Metab. 11 (2), 54-60 (2000) | | | | |
| MEDLINE | 2112178 | | | | |
| PUBMED | 10675891 | | | | |
| REFERENCE | 3 (bases 1 to 11570) | | | | |
| AUTHORS | Yousef,G.M., Luo,L.Y. and Diamandis,E.P. | | | | |
| TITLE | Direct Submision | | | | |
| JOURNAL | Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..11570 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /db_xref="taxon:9606" | | | | |
| | /chromosome="19" | | | | |
| | /map="19q13.3-q13.4" | | | | |
| mrna | join(12101..2231,2210..2293,4762..5023,5763..6019,6105..6238,11092..11570) | | | | |
| | /product="kallikrein-like protein 2 KLK-L2" | | | | |
| | join(2221..2293,4762..5023,5763..6019,6105..6238,11092..11247) | | | | |
| CDS | /codon_start=1 | | | | |
| | /product="kallikrein-like protein 2 KLK-L2" | | | | |

| Query Match | 100.0% | Score 5000 | DB 9 | Length 11570 |
|--|---|--------------|----------|--------------|
| Beet Local Similarity | 100.0% | Pred. No. 0 | | |
| Matches 5000 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| <div> <div>protein_id="AAD36429.1"</div> <div>/db_xref="GI:4589283"</div> <div>/translation="MATARPPEMMVLCALITALLGVTEHYLANNDVSDHPNSTVPSGNSDLGAGGEDARSDSSRIINGSDCMHRTQPMQDALILRNQULVHPQMULTAARCKVFRVRLGHYSLSIPYSESGDMQFOGVKLIHPHVSHPHNDLMLIKNRRIRPTDVRPINSVSHCPSPAGTKCIIVSGGTTKSPVPHVPKYLOCLNTSVLSQKREDAVRQRLDPTMFCAGDKADGRSDSCGDSGGFVVCNGSLQGLVSHMDYFCAPAPNRGVTTLNCKFTKMIQETIQAANS"</div> </div> | | | | |
| 1 | GGGCCCAAGTGAAGGCAAGAGAGAGAGTGGAGACTCCCTCTGCAAAAGTGCTTGAATC | 60 | | |
| 1 | GGGCCCAAGTGAAGGCAAGAGAGAGTGGAGACTCCCTCTGCAAAAGTGCTTGAATC | 60 | | |
| 61 | TCCCTTGCCTTAAATGCAAGGAGAGGAGGCAAGAAACAGGGAAGAGAGGGGTGGGG | 120 | | |
| 61 | TCCCTTGCCTTAAATGCAAGGAGAGGAGGCAAGAAACAGGGAAGAGAGGGGTGGGG | 120 | | |
| 121 | AAGAAAG | 180 | | |
| 121 | AAGAAAG | 180 | | |
| 181 | ACAG | 240 | | |
| 181 | ACAG | 240 | | |
| 241 | ACAAG | 300 | | |
| 241 | ACAAG | 300 | | |
| 301 | AAAGGGGAG | 360 | | |
| 301 | AAAGGGGAG | 360 | | |
| 361 | TTTTTTTTTTTTTTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 420 | | |
| 361 | TTTTTTTTTTTTTTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 420 | | |
| 421 | GATCTCGGCTCAGTCGAGAGCTCGGCTCGGAGTTACAGCCATTCCTCTGCTCAGCTC | 480 | | |
| 421 | GATCTCGGCTCAGTCGAGAGCTCGGCTCGGAGTTACAGCCATTCCTCTGCTCAGCTC | 480 | | |
| 481 | CCAAGTAGCTGGAGCTACAGGCGCGCCCACTACGCGCGCTAATTTTTTGTATTTTTTA | 540 | | |
| 481 | CCAAGTAGCTGGAGCTACAGGCGCGCCCACTACGCGCGCTAATTTTTTGTATTTTTTA | 540 | | |
| 541 | GTAAG | 600 | | |
| 541 | GTAAG | 600 | | |
| 601 | GCCGCGCTCGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAAGCCACGCGCGGCATG | 660 | | |
| 601 | GCCGCGCTCGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAAGCCACGCGCGGCATG | 660 | | |
| 661 | ATCATCTTCTTGAATAGCTGATGTGAACAAGTACCTAAGCCATCAGACTTACCTTTTA | 720 | | |
| 661 | ATCATCTTCTTGAATAGCTGATGTGAACAAGTACCTAAGCCATCAGACTTACCTTTTA | 720 | | |
| 721 | AATATGAGAGTGGGAG | 780 | | |
| 721 | AATATGAGAGTGGGAG | 780 | | |
| 781 | AGAGTGGGAGAGTCACTTGAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 840 | | |
| 781 | AGAGTGGGAGAGTCACTTGAAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 840 | | |
| 841 | CTCTGCTCTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | 900 | | |
| 841 | CTCTGCTCTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | 900 | | |

QY 901 ACCTGTAATCCAGCTATGCTGAGGCTGAGCAGGAGTCACTTGAACCTTGAGGG 960
Db 901 ACCTGTAATCCAGCTATGCTGAGGCTGAGCAGGAGTCACTTGAACCTTGAGGG 960
QY 961 GAGGTTGCAGTGGCCGAGATCAATCAACCCCTCCAGCTGGGCGACAGCAAGACT 1020
Db 961 GAGGTTGCAGTGGCCGAGATCAATCAACCCCTCCAGCTGGGCGACAGCAAGACT 1020
QY 1021 CTGTCTCAATTAATAATAACAAACGAAAGGAGTTGTGTACTTAACTTAATCT 1080
Db 1021 CTGTCTCAATTAATAATAACAAACGAAAGGAGTTGTGTACTTAACTTAATCT 1080
QY 1081 AAAAAAAAAATGCTGTCAACAAATAGACAGAGTAAATAAAGAAAAATTAATGGGCGCA 1140
Db 1081 AAAAAAAAAATGCTGTCAACAAATAGACAGAGTAAATAAAGAAAAATTAATGGGCGCA 1140
QY 1141 AGAATCTTAAGTATATTTTGAACAATCATTCAGAACCTTTAAAAAGAAATCAACA 1200
Db 1141 AGAATCTTAAGTATATTTTGAACAATCATTCAGAACCTTTAAAAAGAAATCAACA 1200
QY 1201 GGCATGAAGAACAGGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
Db 1201 GGCATGAAGAACAGGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
QY 1261 ACNAGGCTCCTTAAGACAGACAGAGAGAGAGAGAGAGTGAAGACAGACAGAG 1320
Db 1261 ACNAGGCTCCTTAAGACAGACAGAGAGAGAGAGAGAGTGAAGACAGACAGAG 1320
QY 1321 AAAAAACAGAGAGAGAGACAGACAGAGAGACAGAGAGGCGAGAGGATGAAGA 1380
Db 1321 AAAAAACAGAGAGAGAGACAGACAGAGAGACAGAGAGGCGAGAGGATGAAGA 1380
QY 1381 GAGAGAGGGGTGAGAGACAGAGATATTTGAGAGACTCGAAGAAATAGCCGAGGA 1440
Db 1381 GAGAGAGGGGTGAGAGACAGAGATATTTGAGAGACTCGAAGAAATAGCCGAGGA 1440
QY 1441 GAAACAAGAGAGATGAAAGAGACTCTGAGAAAAAACAGAGACAAAGTGAAGAG 1500
Db 1441 GAAACAAGAGAGATGAAAGAGACTCTGAGAAAAAACAGAGACAAAGTGAAGAG 1500
QY 1501 AGTATGAGGGGTGAACAGACAGTGTGTGAATGACAAATATGACAGAGAAACAGAA 1560
Db 1501 AGTATGAGGGGTGAACAGACAGTGTGTGAATGACAAATATGACAGAGAAACAGAA 1560
QY 1561 TCCAGGCGCCAAAGATAGTACCAGAGTTGTGAGAAAGCAGATCCTTAAGCTGGGG 1620
Db 1561 TCCAGGCGCCAAAGATAGTACCAGAGTTGTGAGAAAGCAGATCCTTAAGCTGGGG 1620
QY 1621 AGGCAAGGAAAGGGCTGGCTTCCGAGAACCCCTCCATTCCTCCGAGCCAGGAG 1680
Db 1621 AGGCAAGGAAAGGGCTGGCTTCCGAGAACCCCTCCATTCCTCCGAGCCAGGAG 1680
QY 1681 GTTAGGAGTGAATTTCCGAGCTGGGTGGGGGTGTCTGTGGGGGTGAATAGGGGAGCA 1740
Db 1681 GTTAGGAGTGAATTTCCGAGCTGGGTGGGGGTGTCTGTGGGGGTGAATAGGGGAGCA 1740
QY 1741 GAGAGAGCTATTGTAAAGCCGATAGGCACTCATTTGCCCGGAAATGTGCCCAAGGGAG 1800
Db 1741 GAGAGAGCTATTGTAAAGCCGATAGGCACTCATTTGCCCGGAAATGTGCCCAAGGGAG 1800
QY 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCCCAGAGCCCAAGAGAGGAGCTGGCCAGGA 1860
Db 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCCCAGAGCCCAAGAGAGGAGCTGGCCAGGA 1860
QY 1861 AGGCAACAGGCTGAGAAAGTCTGCGGCTGAGCTGGAGCAAAATCCCCACTACCTGG 1920
Db 1861 AGGCAACAGGCTGAGAAAGTCTGCGGCTGAGCTGGAGCAAAATCCCCACTACCTGG 1920
QY 1921 GGGAGAGGGCAATGAGACTGTGTGAGGGGTGCTCAGCAGGCGAGGAAAGAGGTGTCT 1980
Db 1921 GGGAGAGGGCAATGAGACTGTGTGAGGGGTGCTCAGCAGGCGAGGAAAGAGGTGTCT 1980

QY 1981 GTGCGTCTGACCCACACATCTTTCTGTGCCCTCTTCCCTGTCTGAGGCTGTAGA 2040
Db 1981 GTGCGTCTGACCCACACATCTTTCTGTGCCCTCTTCCCTGTCTGAGGCTGTAGA 2040
QY 2041 CTCTATCTTCTGAATTTCTATGTGCTGGGTCTCAGGCAATGCCGATGTGGCCCTGC 2100
Db 2041 CTCTATCTTCTGAATTTCTATGTGCTGGGTCTCAGGCAATGCCGATGTGGCCCTGC 2100
QY 2101 CTGTGTGTTCTCTCTACCTGGGAAATTAAGGTAGGGAGGAGGAGGAGTGGTTAAG 2160
Db 2101 CTGTGTGTTCTCTCTACCTGGGAAATTAAGGTAGGGAGGAGGAGGAGTGGTTAAG 2160
QY 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATCAGGTGACGCGCC 2220
Db 2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATCAGGTGACGCGCC 2220
QY 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTGTGTCTGTATCACAGCTTGTCT 2280
Db 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTGTGTCTGTATCACAGCTTGTCT 2280
QY 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGGTGGAGGGTTGTGGGATTTGGAGGACTGT 2340
Db 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGGTGGAGGGTTGTGGGATTTGGAGGACTGT 2340
QY 2341 CTCTCGGCACTAGAGCGCTGTCTCCCTGGGAACTGTGTAGCCTGGGCACTCGCG 2400
Db 2341 CTCTCGGCACTAGAGCGCTGTCTCCCTGGGAACTGTGTAGCCTGGGCACTCGCG 2400
QY 2401 GACCGGGTGAATGTAGTCTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 GACCGGGTGAATGTAGTCTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY 2461 GACTGCAACGGGT 2520
Db 2461 GACTGCAACGGGT 2520
QY 2521 CAGGTGCACTGACCTTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2521 CAGGTGCACTGACCTTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 2581 ATTTGT 2640
Db 2581 ATTTGT 2640
QY 2641 GTTTTGT 2700
Db 2641 GTTTTGT 2700
QY 2701 TCCCTGAGGTCCCGGATTTGCTGTGCAACAAAAGTGTCACTCACTGAAAGCTGTGACT 2760
Db 2701 TCCCTGAGGTCCCGGATTTGCTGTGCAACAAAAGTGTCACTCACTGAAAGCTGTGACT 2760
QY 2761 GT 2820
Db 2761 GT 2820
QY 2821 TTTTGTGACCGGTGTGACTACTGAACTCTGTGTGAGGGGTGACTGTATGTGACTGTGTG 2880
Db 2821 TTTTGTGACCGGTGTGACTACTGAACTCTGTGTGAGGGGTGACTGTATGTGACTGTGTG 2880
QY 2881 TCTGTGTGAGCGGTGTAAATGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db 2881 TCTGTGTGAGCGGTGTAAATGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
QY 2941 TTTCTGTCTGT 3000
Db 2941 TTTCTGTCTGT 3000
QY 3001 CAGGTGACTGACTTGTCAAGT 3060
Db 3001 CAGGTGACTGACTTGTCAAGT 3060
QY 3061 TGTGACACACCGCATCTGT 3120

Db 3061 TGTGACACACGCGATCTGTGCGTGGACCTGAGACACTGTGATGAGGGGTGTGATCC 3120
Qy 3121 GCTAGGCTCCCGGAGCGGTGTGATCTGTGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180
Db 3121 GCTAGGCTCCCGGAGCGGTGTGATCTGTGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180
Qy 3181 GAGGCAACATGGGCGGTGTGAGACCTGTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3240
Db 3181 GAGGCAACATGGGCGGTGTGAGACCTGTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3240
Qy 3241 GCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3300
Db 3241 GCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3300
Qy 3301 GAGGCAACGCGGCGGTGTGAGACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3360
Db 3301 GAGGCAACGCGGCGGTGTGAGACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3360
Qy 3361 GGGGATACCTGAGCTGCGGAGATGAGGCGAGCTGACCAATGAGAGAACCGGCT 3420
Db 3361 GGGGATACCTGAGCTGCGGAGATGAGGCGAGCTGACCAATGAGAGAACCGGCT 3420
Qy 3421 CTACTAAAAATACAAAAATTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
Db 3421 CTACTAAAAATACAAAAATTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
Qy 3481 GGGAGACTGGGAGGAGAAAAATCGCTTGACCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
Db 3481 GGGAGACTGGGAGGAGAAAAATCGCTTGACCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
Qy 3541 TGTGCGCATGTGCTGTGAGCTGT 3600
Db 3541 TGTGCGCATGTGCTGTGAGCTGT 3600
Qy 3601 GAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
Db 3601 GAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660
Qy 3661 ATGATGT 3720
Db 3661 ATGATGT 3720
Qy 3721 GTTTTATCTGAGAAATTCAGAAATCTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
Db 3721 GTTTTATCTGAGAAATTCAGAAATCTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
Qy 3781 CTGAGCCCTTCTTCTGCTGT 3840
Db 3781 CTGAGCCCTTCTTCTGCTGT 3840
Qy 3841 CCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3900
Db 3841 CCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3900
Qy 3901 CACCCCTTGTGGGCGAGGCTCAACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Db 3901 CACCCCTTGTGGGCGAGGCTCAACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Qy 3961 AGTCAGAGCTTTTGT 4020
Db 3961 AGTCAGAGCTTTTGT 4020
Qy 4021 TGCAGTGGCGT 4080
Db 4021 TGCAGTGGCGT 4080
Qy 4081 GCTTCCACTTCTGAGT 4140
Db 4081 GCTTCCACTTCTGAGT 4140
Qy 4141 GTGTCTTTAGT 4200
Db 4141 GTGTCTTTAGT 4200

Db 4141 GTGTCTTTAGT 4200
Qy 4201 TGCAGT 4260
Db 4201 TGCAGT 4260
Qy 4261 CCCAGCCCAAGT 4320
Db 4261 CCCAGCCCAAGT 4320
Qy 4321 CTACTAGT 4380
Db 4321 CTACTAGT 4380
Qy 4381 TGCAGT 4440
Db 4381 TGCAGT 4440
Qy 4441 GCTCAGCCCGGAGGCTGT 4500
Db 4441 GCTCAGCCCGGAGGCTGT 4500
Qy 4501 TTTACCTGT 4560
Db 4501 TTTACCTGT 4560
Qy 4561 GTTACCTGT 4620
Db 4561 GTTACCTGT 4620
Qy 4621 CTCTTATTTCTGT 4680
Db 4621 CTCTTATTTCTGT 4680
Qy 4681 GGT 4740
Db 4681 GGT 4740
Qy 4741 CAACCAATTTTGT 4800
Db 4741 CAACCAATTTTGT 4800
Qy 4801 CCTCTAACCGGCTGT 4860
Db 4801 CCTCTAACCGGCTGT 4860
Qy 4861 CCCGT 4920
Db 4861 CCCGT 4920
Qy 4921 AGCGT 4980
Db 4921 AGCGT 4980
Qy 4981 TGCATCAGAGT 5040
Db 4981 TGCATCAGAGT 5040

RESULT 5
AC011483/c 107487 bp DNA linear PRI 01-JUL-2002
LOCUS AC011483 Homo sapiens chromosome 19 clone CTB-147c22, complete sequence.
DEFINITION AC011483
AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 107487)
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 1, 2002 this sequence version replaced gi:14971176.
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sbgc.stanford.edu

FEATURES
source
1..107487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C2"
/note="NOTE: Shatter libraries failed to resolve dinucleotide repeat. Unsure number of repeat copies 64998-65494. Forced join 65015."
misc_feature
64998..65494

ORIGIN

Query Match 99.8%; Score 4989; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 1; Gaps 1;
Matches 5000; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGTGGCTTGAGTC 60
DB 60220 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGTGGCTTGAGTC 60161
QY 61 TCCCTGCTTAATATGACGAGAGAGGAGGCAAGAAACAGGAA-GAGGAAGGGGTGGC 119
DB 60160 TCCCTGCTTAATATGACGAGAGAGGAGGAGGCAAGAAACAGGGAAGGAGTGGC 60101
QY 120 GAAGAAG 179
DB 60100 GAAGAAG 60041
QY 180 CACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAAAGAGAGAGAGAGAG 239
DB 60040 CACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAAAGAGAGAGAGAG 59981
QY 240 GACACAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGAGTCCCAAGTACAG 299
DB 59980 GACACAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGAGTCCCAAGTACAG 59921
QY 300 CAAAGGGGACAGAGCAGTTTTCAGGGGTGTATATCATCTCTTTTCTTTTCTTTT 359
DB 59920 CAAAGGGGACAGAGCAGTTTTCAGGGGTGTATATCATCTCTTTTCTTTTCTTTT 59861
QY 360 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 419
DB 59860 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 59801
QY 420 GGAATCTGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCAATCTCTGCTCAGCCT 479
DB 59800 GGAATCTGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCAATCTCTGCTCAGCCT 59741
QY 480 CCCAAGTAGCTGGAGCTACAGGGGCGGCACTAAGCCGCGCTAATTTTGTATTTT 539
DB 59740 CCCAAGTAGCTGGAGCTACAGGGGCGGCACTAAGCCGCGCTAATTTTGTATTTT 59681

QY 540 AGTAGAGACGGGGTTTCAACCGTTTACCGGGAGTGGCTGATCTCTGACCTCGTAGTC 599
DB 59680 AGTAGAGACGGGGTTTCAACCGTTTACCGGGAGTGGCTGATCTCTGACCTCGTAGTC 59621
QY 600 CCGCCGCTCGGCTTCCCAAATGCTGGGATTACAGGGGTGAGCCACCGCGCCGCTCAT 659
DB 59620 CCGCCGCTCGGCTTCCCAAATGCTGGGATTACAGGGGTGAGCCACCGCGCCGCTCAT 59561
QY 660 GATCATCTTCTGACATAGCTGATGACCAAGTACTTAAAGCATCAGCTTACCCCTT 719
DB 59560 GATCATCTTCTGACATAGCTGATGACCAAGTACTTAAAGCATCAGCTTACCCCTT 59501
QY 720 AATATGACAGTTTGGGACAGGACCGTGGCTCATCTGTAAATTCAGACCTTTGGAGG 779
DB 59500 AATATGACAGTTTGGGACAGGACCGTGGCTCATCTGTAAATTCAGACCTTTGGAGG 59441
QY 780 CAGAGTGGGTGAATCACTTAGGCGCAGAGAGTTTGAACACGCTGGCCCAATGCTGAA 839
DB 59440 CAGAGTGGGTGAATCACTTAGGCGCAGAGAGTTTGAACACGCTGGCCCAATGCTGAA 59381
QY 840 ACTGCTCTTTACTTAAAAAATAAAAAATAAAAAATCAAGCGGGGTGTGGGGGCA 899
DB 59380 ACTGCTCTTTACTTAAAAAATAAAAAATAAAAAATCAAGCGGGGTGTGGGGGCA 59321
QY 900 CACTGTATCCAGACTATGCTGAGAGGCTGAGGCAAGAGTCACTTGAACCTTGAAGGC 959
DB 59320 CACTGTATCCAGACTATGCTGAGAGGCTGAGGCAAGAGTCACTTGAACCTTGAAGGC 59261
QY 960 GGAAGTTGAGTGGGCGGAGATCAATACCGCCCTCCAGCTGGGCGCAGAGCAAGAC 1019
DB 59260 GGAAGTTGAGTGGGCGGAGATCAATACCGCCCTCCAGCTGGGCGCAGAGCAAGAC 59201
QY 1020 TCTGCTCAAAATTAATTAATTAACAAAGCAAGAGTGTGTGTAATCTTATATATC 1079
DB 59200 TCTGCTCAAAATTAATTAATTAACAAAGCAAGAGTGTGTGTAATCTTATATATC 59141
QY 1080 TAAAAAATAATGCTGTCAACAAATGAGCAGAAAGTAAATTAAGAAAAATTAATGGGCC 1139
DB 59140 TAAAAAATAATGCTGTCAACAAATGAGCAGAAAGTAAATTAAGAAAAATTAATGGGCC 59081
QY 1140 AAGAACTTAAGTATATTTTGAACAAATCATTCAGAACTTTTAAAAAAGAAATCACAG 1199
DB 59080 AAGAACTTAAGTATATTTTGAACAAATCATTCAGAACTTTTAAAAAAGAAATCACAG 59021
QY 1200 AGGCATAGAGAGACAGGAGAGAAACAGGAGACAGAAACCTGTGGGCCCAAGAGAACAA 1259
DB 59020 AGGCATAGAGAGACAGGAGAGAAACAGGAGACAGAAACCTGTGGGCCCAAGAGAACAA 58961
QY 1260 AACAAAGCTCTTAAGACACAGAGAGAGAGAGAGAGAGTGAAGTGAAGACACAG 1319
DB 58960 AACAAAGCTCTTAAGACACAGAGAGAGAGAGAGAGAGTGAAGTGAAGACACAG 58901
QY 1320 GAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAAGGGATGGAAG 1379
DB 58900 GAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAAGGGATGGAAG 58841
QY 1380 AGAGAGAGGGGTGAGAGAGACAGAGATATTTGAAGAGACTCAGAAAGATAGCCGAGGG 1439
DB 58840 AGAGAGAGGGGTGAGAGAGACAGAGATATTTGAAGAGACTCAGAAAGATAGCCGAGGG 58781
QY 1440 AGAACCAAGAGAGATGAGAGAGACTCTGAGAAAAAACAAGAGCAAAAGATGAGAAAG 1499
DB 58780 AGAACCAAGAGAGATGAGAGAGACTCTGAGAAAAAACAAGAGCAAAAGATGAGAAAG 58721
QY 1500 GAGTATCGAGGGTGAACAGACAGTGTGTGAATGACAAATATGACAGAGAGAGAGAGCA 1559
DB 58720 GAGTATCGAGGGTGAACAGACAGTGTGTGAATGACAAATATGACAGAGAGAGAGCA 58661
QY 1560 ATCCAGGGGCCAAGAAATGTGACCCAGAGTTGATGAGAGAGCAAGATCTTTAAGCTGGGG 1619
DB 58660 ATCCAGGGGCCAAGAAATGTGACCCAGAGTTGATGAGAGAGCAAGATCTTTAAGCTGGGG 58601
QY 1620 GAGGACGAGAAAGGGGCTGGCTGGCTCCGAGAACCTCTCCCATTTCTCCGGGCCAGGG 1679

| | | | |
|----|-------|---|-------|
| Db | 58600 | GAGGCAAGGAAAGGGGCTGGGCTTGGCTTCGAGAAACCCCTCCATTCTCCGGGACAGGGA | 58541 |
| Qy | 1680 | GGTGGGGAAGTGCATTTCCGGAAGTGGGAGGAGGCTTGGGGGTGAGATAGAGGAGAC | 1739 |
| Db | 58540 | GGTGGGGAAGTGCATTTCCGGAAGTGGGAGGAGGCTTGGGGGTGAGATAGAGGAGAC | 58481 |
| Qy | 1740 | AGGAGAGCTAATTTGCTAAGGCCCGGATAGGCACTCATTTGCCGGGAATGTGCCCAAGGGA | 1799 |
| Db | 58480 | AGGAGAGCTAATTTGCTAAGGCCCGGATAGGCACTCATTTGCCGGGAATGTGCCCAAGGGA | 58421 |
| Qy | 1800 | GCAATGGGTGGTTAATCTCAGAGGCCGGGTGCCAGAGCCAGAGGAGGAGCATGTGTCCAGG | 1859 |
| Db | 58420 | GCAATGGGTGGTTAATCTCAGAGGCCGGGTGCCAGAGCCAGAGGAGGAGCATGTGTCCAGG | 58361 |
| Qy | 1860 | AAGGCAAGGCGCTGAGAAAGTCTGGGAGTGAAGCTGGGAGCAAAATCCCCCAACCCCTAACC | 1919 |
| Db | 58360 | AAGGCAAGGCGCTGAGAAAGTCTGGGAGTGAAGCTGGGAGCAAAATCCCCCAACCCCTAACC | 58301 |
| Qy | 1920 | GGGGACAGGGCAAGTGAAGACCTGTGTAGAGGTGAGCTCAGCAGGCAAGGAAAGAGAGGTGTC | 1979 |
| Db | 58300 | GGGGACAGGGCAAGTGAAGACCTGTGTAGAGGTGAGCTCAGCAGGCAAGGAAAGAGAGGTGTC | 58241 |
| Qy | 1980 | TGTGCGGTCTGTGCACCCACATCTTTCTGTGTCCCTCTTGGCCCTGTGTGGAAGGCTGTAG | 2039 |
| Db | 58240 | TGTGCGGTCTGTGCACCCACATCTTTCTGTGTCCCTCTTGGCCCTGTGTGGAAGGCTGTAG | 58181 |
| Qy | 2040 | ACTGCTAATCTTCTTGAAATTCTATATAGTGCCTGGAGTCTCAGGCACTGCGCAATGGTGGCCGT | 2099 |
| Db | 58180 | ACTGCTAATCTTCTTGAAATTCTATATAGTGCCTGGAGTCTCAGGCACTGCGCAATGGTGGCCGT | 58121 |
| Qy | 2100 | CTTGTGTGTTCTCTCTTACCTGTGGGAAATAGGTAGGGGAGGAGGGGAATGTGGTTAAG | 2159 |
| Db | 58120 | CTTGTGTGTTCTCTCTTACCTGTGGGAAATAGGTAGGGGAGGAGGGGAATGTGGTTAAG | 58061 |
| Qy | 2160 | GGCTTCCCCGGATGGCCTTGGGCTTCCAAACCTCTGACATTCGCCATCCAGGTGCAGCGGC | 2219 |
| Db | 58060 | GGCTTCCCCGGATGGCCTTGGGCTTCCAAACCTCTGACATTCGCCATCCAGGTGCAGCGGC | 58001 |
| Qy | 2220 | CATGGCTACAGCAAGAACCCCGCTGGAGTGTGGGTGCTCTGTGCTCTGATCACAGCCCTTGTCT | 2279 |
| Db | 58000 | CATGGCTACAGCAAGAACCCCGCTGGAGTGTGGGTGCTCTGTGCTCTGATCACAGCCCTTGTCT | 57941 |
| Qy | 2280 | TCTTGGGGGTCCACAGGTAAACAGAACTCTGGGAGTGGAGGGATTTGGGATTTGGGAGGACCTG | 2339 |
| Db | 57940 | TCTTGGGGGTCCACAGGTAAACAGAACTCTGGGAGTGGAGGGATTTGGGATTTGGGAGGACCTG | 57881 |
| Qy | 2340 | TCTTGGCGGACCTAGAGCGGCTGTCCCTTGGGGAACTGTGTGAGCCTTGGGCATGACTCCG | 2399 |
| Db | 57880 | TCTTGGCGGACCTAGAGCGGCTGTCCCTTGGGGAACTGTGTGAGCCTTGGGCATGACTCCG | 57821 |
| Qy | 2400 | GGACCGGGGTAAATGTAGTCTCTGTCTGTATCTTTGTGGTTGTGGAGTCGTATATGTGGCCCTG | 2459 |
| Db | 57820 | GGACCGGGGTAAATGTAGTCTCTGTCTGTATCTTTGTGGTTGTGGAGTCGTATATGTGGCCCTG | 57761 |
| Qy | 2460 | TGAATGCAACGAGTGTGTGTGGGGAGGGGAGATGCTTTTCCATTCAGGTGACTGTGTG | 2519 |
| Db | 57760 | TGAATGCAACGAGTGTGTGTGGGGAGGGGAGATGCTTTTCCATTCAGGTGACTGTGTG | 57701 |
| Qy | 2520 | GCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGACTTTAA | 2579 |
| Db | 57700 | GCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGACTTTAA | 57641 |
| Qy | 2580 | GATTGTGTGTGTGCTTCCACAGCTGTGTGGGTGGAAATGCATAGAGCACTGGGGGTTCACATG | 2639 |
| Db | 57640 | GATTGTGTGTGTGCTTCCACAGCTGTGTGGGTGGAAATGCATAGAGCACTGGGGGTTCACATG | 57581 |
| Qy | 2640 | TGT | 2699 |
| Db | 57580 | TGT | 57521 |
| Qy | 2700 | GTCCTTGAAGTCCCGGGAATGTGCTGCAACAAAGTGTCTATCCATGAGAAACTGTGAC | 2759 |

| | | | |
|----|-------|--|-------|
| Db | 57520 | GTCCCTAGAGGTCCCGGGAATTGGCTGCACAAAGTGGTCACTACCAATGGAAAGCTGTGAC | 57461 |
| QY | 2760 | TGTGTGCTGTCTTGCAAGGCGAATTATGTGATTTGGCTTGAGTGTGACCTTATTGATATGCCCT | 2819 |
| Db | 57460 | TGTGTGCTGTCTTGCAAGGCGAATTATGTGATTTGGCTTGAGTGTGACCTTATTGATATGCCCT | 57401 |
| QY | 2820 | ATTTGTGACCGTGTGACTTACCTGAACTCTGTGTAGGGTGACTGTATATGTGACTGTGT | 2879 |
| Db | 57400 | ATTTGTGACCGTGTGACTTACCTGAACTCTGTGTAGGGTGACTGTATATGTGACTGTGT | 57341 |
| QY | 2880 | GTCGTGTGAGGCGCGTGTAAATGCTATCTGATGTGTGATGTGAGCTGTGTGTGAG | 2939 |
| Db | 57340 | GTCGTGTGAGGCGCGTGTAAATGCTATCTGATGTGTGATGTGAGCTGTGTGTGAG | 57281 |
| QY | 2940 | TTTCTGTCTGTGCTTGAGGGAATAGAGGTGCAGGGGTAGCTATCTTGAGGAATGGGTG | 2999 |
| Db | 57280 | TTTCTGTCTGTGCTTGAGGGAATAGAGGGTGCAGGGGTAGCTATCTTGAGGAATGGGTG | 57221 |
| QY | 3000 | CCAGGTGACTGACTTGCAATGTGTGCTGTGTGTGCAGAAAGTATGTGTGCACGTCTGAACAT | 3059 |
| Db | 57220 | CCAGGTGACTGACTTGCAATGTGTGCTGTGTGTGCAGAAAGTATGTGTGCACGTCTGAACAT | 57161 |
| QY | 3060 | CTGTGCAACAACGGGCATCTGTGGGTGGCACTGAGACACTGTGGATAGAGGTGTGCAATCC | 3119 |
| Db | 57160 | CTGTGCAACAACGGGCATCTGTGGGTGGCACTGAGACACTGTGGATAGAGGTGTGCAATCC | 57101 |
| QY | 3120 | CGCTAGGCTGCGCCGGAGAGCTGTGTACTGTGAGACAGAGCTGTATGTTAGCTGCACCTGT | 3179 |
| Db | 57100 | CGCTAGGCTGCGCGGAGAGCGTGTGTACTGTGAGACAGAGCTGTATGTTAGCTGCACCTGT | 57041 |
| QY | 3180 | GGAGGCAACATGGGCGTGTCTGCAGAACTGTGCTGCTTGGCTGTACTGTCTTTGTG | 3239 |
| Db | 57040 | GGAGGCAACATGGGCGTGTCTGCAGAACTGTGCTGCTTGGCTGTACTGTCTTTGTG | 56981 |
| QY | 3240 | CGCGTGGTCTTTGGGGGTGAACTTGTGTAAATGATGTGTGTGCAGGGGCATCAGCAAGGAT | 3299 |
| Db | 56980 | CGCGTGGTCTTTGGGGGTGAACTTGTGTAAATGATGTGTGTGCAGGGGCATCAGCAAGGAT | 56921 |
| QY | 3300 | AGAACCAAGGCGGGGCGGCTGAGCTCAGGCTGTATATCCAGGCGCTTTGGGAGGCGCAGGCG | 3359 |
| Db | 56920 | AGAACCAAGGCGGGGCGGCTGAGCTCAGGCTGTATATCCAGGCGCTTTGGGAGGCGCAGGCG | 56861 |
| QY | 3360 | AGGCGGATCACTGTAGGTGGGAGATGTGAGGCGCAGGCTTGAACCATGTGAGAACCCCTGTC | 3419 |
| Db | 56860 | AGGCGGATCACTGTAGGTGGGAGATGTGAGGCGCAGGCTTGAACCATGTGAGAACCCCTGTC | 56801 |
| QY | 3420 | TCTATCTTAAATATCAAAAATTTATAGCTGTGTGTGTGGCGGTCTGTATATCCAGCTACT | 3479 |
| Db | 56800 | TCTATCTTAAATATCAAAAATTTATAGCTGTGTGTGTGGCGGTCTGTATATCCAGCTACT | 56741 |
| QY | 3480 | CGGAGACTGGGGGCAAAAAATGCGCTTGAACCCGGGAGGTGTGAGGTTTGGGTGAGGCGGAG | 3539 |
| Db | 56740 | CGGAGACTGGGGGCAAAAAATGCGCTTGAACCCGGGAGGTGTGAGGTTTGGGTGAGGCGGAG | 56681 |
| QY | 3540 | ATCGCGGCAATTGCACTTCAGGCTGTGGGCAACAAGGCGAACTTCGCTTGAAAGAAAATA | 3599 |
| Db | 56680 | ATCGCGGCAATTGCACTTCAGGCTGTGGGCAACAAGGCGAACTTCGCTTGAAAGAAAATA | 56621 |
| QY | 3600 | AGAAAAAAAAGGGGTATGAACCAAGTGAATGGGCAAGGAGACTGTATATGATGAGTGGGG | 3659 |
| Db | 56620 | AGAAAAAAAAGGGGTATGAACCAAGTGAATGGGCAAGGAGACTGTATATGATGAGTGGGG | 56561 |
| QY | 3660 | CATCATGTATGTCTGTAGGTCTGTGTGTGTGAGAGAGAAATTGACACGAATTAGAAAGGCA | 3719 |
| Db | 56560 | CATCATGTATGTCTGTAGGTCTGTGTGTGTGAGAGAGAAATTGACACGAATTAGAAAGGCA | 56501 |
| QY | 3720 | TGTTTTCATCTGAGAAATTCAAGAACTTAGGCGTCTTCCCTCCATGTGGCCCCCTTAA | 3779 |
| Db | 56500 | TGTTTTCATCTGAGAAATTCAAGAACTTAGGCGTCTTCCCTCCATGTGGCCCCCTTAA | 56441 |
| QY | 3780 | GCTAGAGCCCTTCTTCTCTGCTCTGCTTCTTGAAACCTTAGCTTCGCGCCATGAGCTCTGAC | 3839 |
| Db | 56440 | GCTAGAGCCCTTCTTCTCTGCTCTGCTTCTTGAAACCTTAGCTTCGCGCCATGAGCTCTGAC | 56381 |

QY 3840 CCCACCTCTTCTCTCAACCAAGCCCTAGGCGAGACTCTAGTGAAGCCCGGCTTAAGGCC 3899
Db 56380 CCCACTCTTCTCTCAACCAAGCCCTAGGCGAGACTCTAGTGAAGCCCGGCTTAAGGCC 56321
QY 3900 ACAACCCCTTGGGCGAGGCTCAACCCCTATTCTGTGGGTAAGCTTCTTAAGAACCCCTTCA 3959
Db 56320 ACAACCCCTTGGGCGAGGCTCAACCCCTATTCTGTGGGTAAGCTTCTTAAGAACCCCTTCA 56261
QY 3960 AAGTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4019
Db 56260 AAGTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 56201
QY 4020 GTGCACTGGGCTGATCTCGGCTCACTGCAACTCTGCTCTCCAGTTCAAGTATCTCG 4079
Db 56200 GTGCACTGGGCTGATCTCGGCTCACTGCAACTCTGCTCTCCAGTTCAAGTATCTCG 56141
QY 4080 TGGCTTCACTCTCTGAGTGAAGTGGGATTTACAGGTGGCGCCACACGCTGGCTAAATTTT 4139
Db 56140 TGGCTTCACTCTCTGAGTGAAGTGGGATTTACAGGTGGCGCCACACGCTGGCTAAATTTT 56081
QY 4140 TGTGTCTTTAGTGAAGAGGCTTTTCACTTTTGGCGAGGCTGTCTAAACTCTCCAC 4199
Db 56080 TGTGTCTTTAGTGAAGAGGCTTTTCACTTTTGGCGAGGCTGTCTAAACTCTCCAC 56021
QY 4200 CTGAGGTGATCCGCCCACTCGGCTCCAGAGTGGGGTTACAGGCGTGAAGCACC 4259
Db 56020 CTGAGGTGATCCGCCCACTCGGCTCCAGAGTGGGGTTACAGGCGTGAAGCACC 55961
QY 4260 CCCCCAGCCCAAGTCAAGAGCTCTTTATAGAGAGCTCTTAACATGTAACTGACCTGCG 4319
Db 55960 CCCCCAGCCCAAGTCAAGAGCTCTTTATAGAGAGCTCTTAACATGTAACTGACCTGCG 55901
QY 4320 CCTAATGATGAATTCGAAACCCCTCTGCTCTGAGCTTCAAGCCCACTCACTCAAGGC 4379
Db 55900 CCTAATGATGAATTCGAAACCCCTCTGCTCTGAGCTTCAAGCCCACTCACTCAAGGC 55841
QY 4380 CTGACCCCACTTCTTGAAGCAAGTTCATCTTAAAGCCCTGCTCTCCATCCCA 4439
Db 55840 CTGACCCCACTTCTTGAAGCAAGTTCATCTTAAAGCCCTGCTCTCCATCCCA 55781
QY 4440 GGGCTCAGGCCCCCAAGCTTTGGCACTACCCCTGAGCTTGTCAAGAAATCTGTACCA 4499
Db 55780 GGGCTCAGGCCCCCAAGCTTTGGCACTACCCCTGAGCTTGTCAAGAAATCTGTACCA 55721
QY 4500 TTTTACCTCAAGATGATTTCTAGCAATTTCCAGAAATCTGAGGTTCAAGTTC 4559
Db 55720 TTTTACCTCAAGATGATTTCTAGCAATTTCCAGAAATCTGAGGTTCAAGTTC 55661
QY 4560 AGTAACTTCACTGAGCTTGGGCTCTGCTTGTAGCTTGAAGCTTGAAGGTTGCC 4619
Db 55660 AGTAACTTCACTGAGCTTGGGCTCTGCTTGTAGCTTGAAGCTTGAAGGTTGCC 55601
QY 4620 ACTCTTATTTCTCAAGGCTTGGCCCTGCTGCTGCTCAAGATGTCAAGCACTTCTAGC 4679
Db 55600 ACTCTTATTTCTCAAGGCTTGGCCCTGCTGCTGCTCAAGATGTCAAGCACTTCTAGC 55541
QY 4680 TGGTCTGGGCTCTTGAAGTGAAGCAACCCCGAGCCCAAGCCCGCTCTGAGCCCGC 4739
Db 55540 TGGTCTGGGCTCTTGAAGTGAAGCAACCCCGAGCCCAAGCCCGCTCTGAGCCCGC 55481
QY 4740 CCAACCATTTTCCGTTTCCAGAGAGATTTCTGCAACATGATGTTTCTTGAACAC 4799
Db 55480 CCAACCATTTTCCGTTTCCAGAGAGATTTCTGCAACATGATGTTTCTTGAACAC 55421
QY 4800 CCTCTTAACAACCGTGCCTCTTGGAGCAACAGAGACTTGGAGCTTGGGCGCGGGAAGAC 4859
Db 55420 CCTCTTAACAACCGTGCCTCTTGGAGCAACAGAGACTTGGAGCTTGGGCGCGGGAAGAC 55361
QY 4860 GCGCGGCTGGAGTACAGAGAGAGCCGATATTAAGGATCCGATCGCATATGCAACAC 4919
Db 55360 GCGCGGCTGGAGTACAGAGAGAGCCGATATTAAGGATCCGATCGCATATGCAACAC 55301

QY 4920 CAGCCGTGCGAGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTG 4979
Db 55300 CAGCCGTGCGAGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTG 55241
QY 4980 GTGCATCCACAGTGGCTGCTC 5000
Db 55240 GTGCATCCACAGTGGCTGCTC 55220
RESULT 6
AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
LOCUS Homo sapiens serine protease gene cluster, complete sequence.
DEFINITION AF243527
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL Gene 257 (1), 119-130 (2000)
MEDLINE 20510030
PubMed 11054574
REFERENCE
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Direct Submision
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
source location/Qualifiers
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/complement(<6963..>11517)
/gene="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCIALSLGTAAPPIQSHIVGMECEHSPQWAAALYH
FSTFOGGLIVHROWLTAALHCISDNYOLMGRHLFPDENTAOAVHVESPPHGFN
MSLENNHROADEVSHDMLRLTREPADTTIDAVKVELEPPEPVSTCLASMGSG
IEPEPSPFDLQCYDLKILPDECKRAHVQRTDPMACVGHLEBGSKDTCVDSGGL
MCDGVLAQVTSAGVPCGTPNKPVSVAVLSVKXIEPTIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
/product="ACO protease"
complement(join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
/note="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MMILTLTSLFLASTPAADGDKLFGDECAPIHSQPMQVALYERGR
FNCGASLISPHVIVLSAAHQSRFMRVLRGEHLNRKRDPEQLRTTSRYIPIHRYEARS
HRNDIMLRIVQPARLNPQVRPAVLPTRCPHGEACVSGMGLVLSVPSVSLPDTLHC

ANISIDTSCKSKYSGRLNTNWCAGAEGRGASCEGDSGGLVCGGIIQGIYVSGD
VPCDNTKPGVYTVCHYLEMIRETMRN"
/gene="KLK3"
/join(<42595, .42640, 43880, .44039, 45669, .45955, 46099, .46235,
47614, .>47769)
/gene="KLK3"
/product="prostate specific antigen"
/join(42595, 42640, 43880, .44039, 45669, .45955, 46099, .46235,
47614, .47769)
/gene="KLK3"
/note="serine protease"
/codon_start=1
/product="prostate specific antigen"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MWPVVFELTSTVIGAAPILISRIVGMECEKHSOPWQVAVS
RGRVACGVVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK2"
/join(<61139, .61184, 62391, .62550, 64142, .64428, 64542, .64678,
66074, .>66229)
/gene="KLK2"
/product="glutular kallikrein"
/join(61139, 61184, 62391, .62550, 64142, .64428, 64542, .64678,
66074, .66229)
/gene="KLK2"
/note="serine protease"
/codon_start=1
/product="glutular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="IMDLVLSIALSVCTGAVPIQSRIVGMECEKHSOPWQVAVS
HGMACGVVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK2"
/join(<94602, .94754, 96027, .96163, 96247, .96497,
96919, .97081, 98345, .>98405)
/product="serine protease"
/complement(join(94602, 94754, 96027, .96163, 96247, .96497,
96919, .97081, 98345, .98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATRGNPMGWFLLGYLLGVASGLVSGSQIITNGEDCSHSPQW
QALVWNEILFGSGVTVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK5"
/complement(join(131301, .131456, 136310, .136443,
136529, .136785, 137525, .137690, 140255, .>140303))
/note="synonym: SCTE"
/complement(join(<131301, .131456, 136310, .136443,
136529, .136785, 137525, .137690, 140255, .>140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWTLCALITALLIGTVDARSDDSSRLTNGSDCDHATQWQAL
LIRPNQLVCGVTVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK5"
/complement(join(131301, .131456, 136310, .136443,
136529, .136785, 137525, .137690, 140255, .>140303))
/note="synonym: SCTE"
/complement(join(<131301, .131456, 136310, .136443,
136529, .136785, 137525, .137690, 140255, .>140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWTLCALITALLIGTVDARSDDSSRLTNGSDCDHATQWQAL
LIRPNQLVCGVTVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"

mRNA
complement(join(<146834, .146986, 149628, .149764,
151186, .151433, 155052, .155208, 155948, .>155987))
/product="protease M"
/complement(join(146834, .146986, 149628, .149764,
151186, .151433, 155052, .155208, 155948, .>155987))
/note="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKRLVNLVSLAANAEEQNKLVHGPPCDKTSHPYQALVTSGH
LLCGVLIHPVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK7"
/complement(join(<165420, .165575, 167672, .167808,
168124, .168371, 169651, .169798, 170211, .>170283))
/note="synonym: SCCE"
/complement(join(<165420, .165575, 167672, .167808,
168124, .168371, 169651, .169798, 170211, .>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
/complement(join(165420, .165575, 167672, .167808,
168124, .168371, 169651, .169798, 170211, .>170283))
/note="serine protease"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLLPLQIILSLAETAGEAOGDKIIDGAPCARSGHPQ
VLLSGNOLHCGSVTVNERVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK7"
/complement(join(<183943, .184098, 185635, .185768,
187865, .188127, 188293, .188452, 188967, .>189036))
/product="neutropsin"
/complement(join(183943, .184098, 185635, .185768,
187865, .188127, 188293, .188452, 188967, .>189036))
/note="serine protease; also called ovastin"
/codon_start=1
/product="neutropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRRAKTMFLLIGAGNAGSHADEKTLGHEGCPHSQ
PMQALFQGGQLCGVTVHPQWVLTAAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"
/gene="KLK7"
/complement(join(<190980, .191129, 191573, .191709,
194324, .194589, 197048, .197204, 197370, .>197412))
/product="kallikrein-like 3"
/complement(join(190980, .191129, 191573, .191709,
194324, .194589, 197048, .197204, 197370, .>197412))
/note="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MGLGLCALSLLAGNADTAIGAEBRCPNSOPQALFHLT
RLPGATLSDPWLTAHICENKSVILGRSLFHPEDTQVPSVSPHPLVD
MSLLKNFLRPDSDSHDMLRLSEPAELDAVKMDLPTQEPALGTTGCVASGMSI
EPPEFLTPKLCQCDLHVISNDVCAQVHPQKVKPMICAGRWGKSTCGSDSGPLV
CNGVLQGITSMGSPCALPERPSLYTVHVRKMKITIVANP"

Query Match 99.84; Score 4989; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 142548 GGGCCGAGTGAAGGCAAGAGAGTGAAGCTCCCTCTGCAAAAGTGGCTTGAATC 60
1 GGGCCGAGTGAAGGCAAGAGAGTGAAGCTCCCTCTGCAAAAGTGGCTTGAATC 60
GGGCCGAGTGAAGGCAAGAGAGTGAAGCTCCCTCTGCAAAAGTGGCTTGAATC 142489

| | | | |
|---|--------|--|--------|
| D | 140328 | CATGGCTACAGCAAGACCCCTTGGATGTGGTCTCTGTCTGTATCAAGCTTGTCT | 140269 |
| Q | 2280 | TTTGGGGGTCA CAGGTAA CAGAA CTCTGGGGTGGAGGGTGTGGATTTGGAGACTG | 2339 |
| D | 140268 | TTTGGGGGTCA CAGGTAA CAGAA CTCTGGGGTGGAGGGTGTGGATTTGGAGACTG | 140209 |
| Q | 2340 | TTTCTGGGGCACTAGACCGCTGTCCCTGGGGGACGTGTGACCTTGGGATATACCTCC | 2339 |
| D | 140208 | TTCTGTGGGCACTAGACCGCTGTCCCTGGGGGACGTGTGACCTTGGGATATACCTCC | 140149 |
| Q | 2400 | GACCCGGGTGATGTGAGTCTCTGTCTGTACTTGTGGTGTGCATGTATGTGGCCTG | 2459 |
| D | 140148 | GACCCGGGTGATGTGAGTCTCTGTCTGTACTTGTGGTGTGCATGTATGTGGCCTG | 140089 |
| Q | 2460 | TGACTGCCACGGGTGTGTGTGGGGAGGGGGATGCTTTTCCCATATCAGTGA CTGTGC | 2519 |
| D | 140088 | TGACTGCCACGGGTGTGTGTGGGGAGGGGGATGCTTTTCCCATATCAGTGA CTGTGC | 140029 |
| Q | 2520 | GCAAGTGGCACTGACCCCTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 2579 |
| D | 140028 | GCAAGTGGCACTGACCCCTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 139969 |
| Q | 2580 | GATTGT | 2639 |
| D | 139968 | GATTGT | 139909 |
| Q | 2640 | TGT | 2699 |
| D | 139908 | TGT | 139849 |
| Q | 2700 | GTCTCTGACGT | 2759 |
| D | 139848 | GTCTCTGACGT | 139789 |
| Q | 2760 | TGT | 2819 |
| D | 139788 | TGT | 139729 |
| Q | 2820 | ATTTGTGACCGGT | 2879 |
| D | 139728 | ATTTGTGACCGGT | 139669 |
| Q | 2880 | GTCTGT | 2939 |
| D | 139668 | GTCTGT | 139609 |
| Q | 2940 | TTTCTGT | 2999 |
| D | 139608 | TTTCTGT | 139549 |
| Q | 3000 | CCAGGTGACCTGT | 3059 |
| D | 139548 | CCAGGTGACCTGT | 139489 |
| Q | 3060 | CTGT | 3119 |
| D | 139488 | CTGT | 139429 |
| Q | 3120 | CGCTAGGCTGT | 3179 |
| D | 139428 | CGCTAGGCTGT | 139369 |
| Q | 3180 | GAGGCAACATGT | 3239 |
| D | 139368 | GAGGCAACATGT | 139309 |
| Q | 3240 | CGCTGT | 3299 |
| D | 139308 | CGCTGT | 139249 |
| Q | 3300 | AGAACAGGCGGGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3359 |
| D | 139248 | AGAACAGGCGGGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 139189 |
| Q | 3360 | AGGCGGATTCATCTGAGGTGGGAGATTCGAGGCCAGCTGTGACCAATGTAGAAACCCG | 3419 |
| D | 139188 | AGGCGGATTCATCTGAGGTGGGAGATTCGAGGCCAGCTGTGACCAATGTAGAAACCCG | 139129 |
| Q | 3420 | TTTACTAAAAATTA CAAAAA TTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3479 |
| D | 139128 | TTTACTAAAAATTA CAAAAA TTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 139069 |
| Q | 3480 | CGGAGACTGGGGCAGAAAAATGCTTGAACCCGGGAGGTGAGGTGGGTGGAGCCGAG | 3539 |
| D | 139068 | CGGAGACTGGGGCAGAAAAATGCTTGAACCCGGGAGGTGAGGTGGGTGGAGCCGAG | 139009 |
| Q | 3540 | ATGCGGCAATTTGCACTTCAGCTTGGGCAACAAGAGGAACTCCGTCTCCAAAGAAAA | 3599 |
| D | 139008 | ATGCGGCAATTTGCACTTCAGCTTGGGCAACAAGAGGAACTCCGTCTCCAAAGAAAA | 138949 |
| Q | 3600 | AGAAAAAAGGATTAAGCAACAGTGAATGGGCA CGGAGAGACTGATGTGAGTGGGG | 3659 |
| D | 138948 | AGAAAAAAGGATTAAGCAACAGTGAATGGGCA CGGAGAGACTGATGTGAGTGGGG | 138889 |
| Q | 3660 | CATGATGT | 3719 |
| D | 138888 | CATGATGT | 138829 |
| Q | 3720 | TGTTTTATTTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGCCCTTA | 3779 |
| D | 138828 | TGTTTTATTTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGCCCTTA | 138769 |
| Q | 3780 | GCTGAGCCCTTCTTCTTCTGCTGCTGCTTGTGGAACCTAGCTCCGACATGAGCTGAC | 3839 |
| D | 138768 | GCTGAGCCCTTCTTCTTCTGCTGCTGCTTGTGGAACCTAGCTCCGACATGAGCTGAC | 138709 |
| Q | 3840 | CCCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 3899 |
| D | 138708 | CCCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 138649 |
| Q | 3900 | ACACCCCTTTGGGCGAGGCTTCCACCCCTTATTTGTGGGATCACTTCTTGAACCCCTT | 3959 |
| D | 138648 | ACACCCCTTTGGGCGAGGCTTCCACCCCTTATTTGTGGGATCACTTCTTGAACCCCTT | 138589 |
| Q | 3960 | AACTGAGAGCTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGCTTGTCTCTTCCAGGCTGA | 4019 |
| D | 138588 | AACTGAGAGCTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGCTTGTCTCTTCCAGGCTGA | 138529 |
| Q | 4020 | GTGAGGTGGGTATCTCGGCTCACTGCAACCTCTGCTTCCAGGTTCAAGTATTCG | 4079 |
| D | 138528 | GTGAGGTGGGTATCTCGGCTCACTGCAACCTCTGCTTCCAGGTTCAAGTATTCG | 138469 |
| Q | 4080 | TGCTTCACTCTCTGAGTGTGGATTA CAGGTGTGGCGCCACACCGCTTGGCTAATTTT | 4139 |
| D | 138468 | TGCTTCACTCTCTGAGTGTGGATTA CAGGTGTGGCGCCACACCGCTTGGCTAATTTT | 138409 |
| Q | 4140 | TGTGTCTTTAGTGAAGACAGGGTTTTCACTTTTGTGGCAGAGCTGTCTCAACCTCCAC | 4199 |
| D | 138408 | TGTGTCTTTAGTGAAGACAGGGTTTTCACTTTTGTGGCAGAGCTGTCTCAACCTCCAC | 138349 |
| Q | 4200 | CTCAGGTATCCGCCCACTCGGCTCCAGAGTGTGTGGGTTTACAGGCTGTAGCCACG | 4259 |
| D | 138348 | CTCAGGTATCCGCCCACTCGGCTCCAGAGTGTGTGGGTTTACAGGCTGTAGCCACG | 138289 |
| Q | 4260 | CCCCCAGCCAAAGTCAAGACTCTTTATAGAGACTTAA CAGTGAACCTGTACCTTGGC | 4319 |
| D | 138288 | CCCCCAGCCAAAGTCAAGACTCTTTATAGAGACTTAA CAGTGAACCTGTACCTTGGC | 138229 |
| Q | 4320 | CCTAATTAAGTCAATTCAAACCCCTTCCGAGCTTCAAGCCCTGACCCCACTGAGGC | 4379 |
| D | 138228 | CCTAATTAAGTCAATTCAAACCCCTTCCGAGCTTCAAGCCCTGACCCCACTGAGGC | 138169 |
| Q | 4380 | CTGACCCCACTTCTTGAAGCAAGTTCATCTTAAAGCCCTGTGTCTTCCATCCCA | 4439 |
| D | 138168 | CTGACCCCACTTCTTGAAGCAAGTTCATCTTAAAGCCCTGTGTCTTCCATCCCA | 138109 |

| QY | 4440 | GGCTCCAGCCCCCAGACGTTTGACATACCCGAGCTTGTCAGAGAAATCTGTACCCAA | 4499 |
|------------|--|---|---------|
| Db | 138108 | GGCTCCAGCCCCCAGACGTTTGACATACCCGAGCTTGTCAGAGAAATCTGTACCCAA | 138049 |
| QY | 4500 | TTTAAACCTGACATGATGTTCTAGCCCAATTCAGAAATCTGTAGAGTCCAGTGAAGTCC | 4559 |
| Db | 138048 | TTTAAACCTGACATGATGTTCTAGCCCAATTCAGAAATCTGTAGAGTCCAGTGAAGTCC | 1379889 |
| QY | 4560 | AGTAAACCTGACATGAGCCCTGGGCTGTCTCTTGAAGTTAGAGCTTGAGAGAGTGC | 4619 |
| Db | 137988 | AGTAAACCTGACATGAGCCCTGGGCTGTCTCTTGAAGTTAGAGCTTGAGAGAGTGC | 137929 |
| QY | 4620 | ACTCTTAATCTCAGAGCCCTGGCCCTCGCCCTCAGCATGTGAGACCAACCCCTTAGC | 4679 |
| Db | 137928 | ACTCTTAATCTCAGAGCCCTGGCCCTCGCCCTCAGCATGTGAGACCAACCCCTTAGC | 137865 |
| QY | 4660 | TGCTCTGAGCTCTTGAAGTCTGAAACCCACCCCAAGCCCGCTCTGAGCCCGC | 4739 |
| Db | 137868 | TGCTCTGAGCTCTTGAAGTCTGAAACCCACCCCAAGCCCGCTCTGAGCCCGC | 137809 |
| QY | 4740 | CCAACCCANTTTCGCTTCCAGAGCAATGTTCTCCAAATGATATTTCTCTGACAC | 4799 |
| Db | 137808 | CCAACCCANTTTCGCTTCCAGAGCAATGTTCTCCAAATGATATTTCTCTGACAC | 137749 |
| QY | 4800 | CCCTCTAAACACCGTCGCCCTCTGGAGCAACCAAGACCTTGGAGCTGGAGCGGAGAGAC | 4859 |
| Db | 137748 | CCCTCTAAACACCGTCGCCCTCTGGAGCAACCAAGACCTTGGAGCTGGAGCGGAGAGAC | 137689 |
| QY | 4860 | GCCCGGTGGATGACAGCAGCAGCCGCGATCATGATGATATCCGATGCGATATGACACC | 4919 |
| Db | 137688 | GCCCGGTGGATGACAGCAGCAGCCGCGATCATGATGATATCCGATGCGATATGACACC | 137629 |
| QY | 4920 | CAGCGGTGGACAGCGCGCTGTGCTAAAGGCCCAACACACTCTGATCGCGGCGGAGTGG | 4979 |
| Db | 137628 | CAGCGGTGGACAGCGCGCTGTGCTAAAGGCCCAACACACTCTGATCGCGGCGGAGTGG | 137569 |
| QY | 4980 | GTGATTCACACAGTGGCTGCTC 5000 | |
| Db | 137568 | GTGATTCACACAGTGGCTGCTC 137548 | |
| RESULT 7 | | | |
| AC027602/c | | | |
| LOCUS | 217346 bp | DNA | linear |
| DEFINITION | Hom sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT | | |
| SEQUENCE | 12 unordered pieces. | | |
| AC027602.4 | GI:11178143 | | |
| VERSION | | | |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT. | | |
| SOURCE | Hom sapiens (human) | | |
| ORGANISM | Hom sapiens | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| REFERENCE | | | |
| AUTHORS | | | |

| | | |
|---|---------|-----------|
| TITLE | JOURNAL | REFERENCE |
| AUTHORS | | |
| <p>O'Neill,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisaní,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Seery,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliév,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.</p> | | |
| Direct Submission | | |
| Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | |
| 3 (bases 1 to 217346) | | |
| <p>Blyden,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Borkgatter,B., Brown,A., Burkett,G., Camporiano,A., Castle,A., Choehl,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dearlino,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Iglew,I., Johnson,R., Jones,C., Kam,L., Karatas,A., Labocque,K., Lamaras,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrum,J., Menue,L., Mhova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisaní,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Seery,P., Sougniez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliév,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.</p> | | |
| Direct Submission | | |
| Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | |
| On Nov 15, 2000 this sequence version replaced gi:11136831. | | |
| All repeats were identified using RepeatMasker: | | |
| Smit, A.P.A. & Green, P. (1996-1997) | | |
| http://ftp.genome.washington.edu/RM/RepeatMasker.html | | |
| ----- Genome Center | | |
| Center: Whitehead Institute/ MIT Center for Genome Research | | |
| Center code: WIBR | | |
| Web site: http://www-seq.wi.mit.edu | | |
| Contact: sequence_submissions@genome.wi.mit.edu | | |
| ----- Project Information | | |
| Center project name: L916 | | |
| Center clone name: 795_B6 | | |
| ----- Summary Statistics | | |
| Sequencing vector: M13: M77815; 31% of reads | | |
| Sequencing vector: Plasmid; n/a; 69% of reads | | |
| Chemistry: Dye-terminator Big Dye; 100% of reads | | |
| Assembly program: Phrap; version 0.960731 | | |
| Consensus quality: 210748 bases at least Q40 | | |
| Consensus quality: 21655 bases at least Q30 | | |
| Consensus quality: 215058 bases at least Q20 | | |
| Insert size: 19400; agarose-fp | | |
| Insert size: 216246; sum-of-contigs | | |
| Quality coverage: 11.9 in Q20 bases; agarose-fp | | |
| Quality coverage: 10.7 in Q20. | | |
| NOTE: This is a "working draft" sequence. It currently | | |
| consists of 12 contigs. The true order of the pieces | | |
| is not known and their order in this sequence record is | | |
| arbitrary. Gaps between the contigs are represented as | | |
| runs of N, but the exact sizes of the gaps are unknown. | | |
| This record will be updated with the finished sequence | | |
| as soon as it is available and the accession number will | | |
| be preserved. | | |
| 1 | | |
| 8149: contig of 8149 bp in length | | |
| 8150 | | |
| 8249: gap of 100 bp | | |
| 8250 | | |
| 9592: contig of 1343 bp in length | | |
| 8250 | | |
| 9692: gap of 100 bp | | |
| 9693 | | |
| 10733: contig of 1041 bp in length | | |
| 10734 | | |
| 10833: gap of 100 bp | | |
| 13519: contig of 2686 bp in length | | |

| FEATURES | |
|----------|---|
| * | 13520 13619: gap of 100 bp |
| * | 13620 17510: contig of 3891 bp in length |
| * | 17511 17610: gap of 100 bp |
| * | 17611 24602: contig of 6992 bp in length |
| * | 24603 24702: gap of 100 bp |
| * | 35434 35434: contig of 10732 bp in length |
| * | 35435 35534: gap of 100 bp |
| * | 35535 124474: contig of 88940 bp in length |
| * | 124475 124574: gap of 100 bp |
| * | 124575 134664: contig of 10090 bp in length |
| * | 134665 134764: gap of 100 bp |
| * | 134765 162343: contig of 27579 bp in length |
| * | 162344 162344: gap of 100 bp |
| * | 162444 208817: contig of 4674 bp in length |
| * | 208917 209017: gap of 100 bp |
| * | 208918 217946: contig of 8359 bp in length. |
| * | 209018 Location/Qualifiers |

ORIGIN

| | | | | |
|----------------------|--------------|--------------|---------------|---------------|
| Query Match | 84.9% | Score 4243.4 | DB 2 | Length 217346 |
| Best Local Similarly | 99.0% | Pred. No. 0 | | |
| Matches 4335 | Conservative | 0 | Mismatches 11 | Indels 35 |
| | | | | Gaps 5 |

| QY | 654 | GGCCATGATCATCTTTCTTGACTATGCTGATGTGACACAGTAACCTAAAGCCATCAGACTTA | 713 |
|----|--------|--|--------|
| Db | 203789 | GTCATATGATCATCTTCTTGACTATGCTGATGTGACACAGTAACCTAAAGCCATCAGACTTA | 203730 |
| QY | 714 | CCCTTTAAATATGCACTTTGGGCGCAGGCAACCGTGCTCATGCTGTAAATTCACGCACTTT | 773 |
| Db | 203729 | CCCTTTAAATATGCACTTTGGGCGCAGGCAACCGTGCTCATGCTGTAAATTCACGCACTTT | 203670 |
| QY | 774 | GGGAGGCGAGAGGTGGGTGAAATCACTTTGAGGCCAGAGAGTTTGAGAACAGAGCTGGCCAAAT | 833 |
| Db | 203669 | GGGAGGCGAGAGGTGGGTGAAATCACTTTGAGGCCAGAGAGTTTGAGAACAGAGCTGGCCAAAT | 203610 |
| QY | 834 | GGTGAACCTCTGTCTTTTACTAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCGT | 893 |

| | | | |
|----|--------|---|---------|
| QY | 1970 | CAGAGGTGCTGTGGTCCCTGCACCCCAKATCTTTCTGTGCCCTCCTTGCCCTGTCTGG | 2029 |
| Db | 202470 | GAGAGGTGTCTGTGCTCTGTGCACCACAATCTTTCTGTGCCCTCCTTGCCCTGTCTGG | 2024111 |
| QY | 2030 | AGCGTCGTGAATCTCATCTTTCTGAATTCTAATGTACTCTGTGGCTCTCAGCGCAGTGCAGT | 2089 |
| Db | 202410 | AGCGTCGTGAATCTCATCTTTCTGAATTCTAATGTACTCTGTGGCTCTCAGCGCAGTGCAGT | 2023511 |
| QY | 2090 | GGTGGCCCGTCTTGTGTGTTCTCTCTTACCTGGGGAAATPAGGTAGGGAGGGAGAA | 2149 |
| Db | 202350 | GGTGGCCCGTCTTGTGTGTTCTCTCTTACCTGGGGAAATCAGGTAGGGAGGGAGGGAA | 2022911 |
| QY | 2150 | GTGGGTTTAAGGCTCCCCGGAATGACTGTGGAGCTCCCAACCCCTGTGAATTTCCCATTCCAG | 2209 |
| Db | 202290 | GTGGGTTTAAGGCTCCCCGGAATGACTGTGGAGCTCCCAACCCCTGTGAATTTCCCATTCCAG | 2022111 |
| QY | 2210 | GTGCACGGCCATGTGCTACAGCAAGA(C)CCCCCTGATGTGGGTCTGTGTCTGTATCA | 2269 |
| Db | 202230 | GTGCACGGCCCATGTGCTACAGCAAG(C)CCCCCTGATGTGGGTCTGTGTCTGTATCA | 2021711 |
| QY | 2270 | CAGCCTTGCTTGTGGGGGTCA CAGGTAACCAAG(A)CTGTGGGGTAGAGGTTGTGGGATT | 2329 |
| Db | 202170 | CAGCCTTGCTTGTGGGGGTCA CAGGTAACCAAG(A)CTGTGGGGTAGAGGTTGTGGGATT | 2021111 |
| QY | 2330 | GGGAGGACTGTCTCTGCGCACCTAGAGCCCTGTCCCTGGGGAACTGTGTGACCTGGG | 2389 |
| Db | 202110 | GGGAGGACTGTCTCTGCGCACCTAGAGCCCTGTCCCTGGGGAACTGTGTGAGCTGGG | 2020511 |
| QY | 2390 | CATGACTCCGGGACCGGGTGAAATGTGAATCTGTCTGTATCTTGTGGTGTGTGCATCTGA | 2449 |
| Db | 202050 | CATGACTCCGGGACCGGGTGAAATGTGAATCTGTCTGTATCTTGTGGTGTGTGCATCTGA | 2019911 |
| QY | 2450 | TGTGACCCCTGTACTGCCACGGGTGTGTCTGGGGAGGGGGAATGCCTTTTCCCATTCAGG | 2509 |
| Db | 201990 | TGTGACCCCTGTACTGCCACGGGTGTGTCTGGGGAGGGGGAATGCCTTTTCCCATTCAGG | 2019311 |
| QY | 2510 | TGACTGTGCGGACGTTGSCACTG(A)CCCTTTGAGGCTGTGTGTGGTTTTGTGATTGTGT | 2569 |
| Db | 201930 | TGACTGTGCGGACGTTGSCACTG(A)CCCTTTGAGGCTGTGTGTGGTTTTGTGATTGTGT | 2018711 |
| QY | 2570 | GTGCATTTTAAGATTTGTGTGGCTCCA CAGCTGTGTGGGTAATGTGATPAGCACTGGGG | 2629 |
| Db | 201870 | GTGCATTTTAAGATTTGTGTGGCTCCA CAGCTGTGTGGGTAATGTGATPAGCACTGGGG | 2018111 |
| QY | 2630 | GTGTTCACTGTGTGTTTGGCTGTGTGTGTGTGACTTGGCATTTGATATGACTGCAG----- | 2684 |
| Db | 201810 | GTGTTCACTGTGTGTTTGGCTGTGTGTGTGTGACTTGGCATTTGATATGACTGCAGTCAAT | 2017511 |
| QY | 2685 | -----GTATCTGCAG-----TTCTGTCTCTGAGGTCCTCCGGATTTG | 2720 |
| Db | 201750 | ATAACATATCCGTATGTATGTATGACAGATACATATACATATCCCTGAGGTCCTCCGGATTTG | 2016911 |
| QY | 2721 | CGTGCACAAAGATGGTCACTACACATGAGAAAGCTGTGACTGTGTGCTGTGACGCGAT | 2780 |
| Db | 201690 | CGTGCACAAAGATGGTCACTACACGCGGAAAGCTGTGACTGTGTGCTGTGACGCGAT | 2016311 |
| QY | 2781 | TATGTGATTTGGCTGTGAGTGTGACGTTATGTAGATCCCGTATTTGTGACCGGTGTGACTACC | 2840 |
| Db | 201630 | TATGTGATTTGGCTGTGAGTGTGACGTTATGTAGATCCCGTATTTGTGACCGGTGTGACTACC | 2015711 |
| QY | 2841 | TGAAGCTCTGTGTAGGGGTGACGTATGTGAC(T)GTGTGTGTCTGTGTGTGTGTGTGTGTAAA | 2900 |
| Db | 201570 | TGAAGCTCTGTGTAGGGGTGACGTATGTGAC(T)GTGTGTGTGTGTGTGTGTGTGTGTGTAAA | 2015111 |
| QY | 2901 | TGCTACTGTATGTGTGATGTGTGACGCTGTGTGTCTGTGAGATTTCTGTCTCTGCCTGAGGG | 2960 |
| Db | 201510 | TGCTACTGTATGTGTGATGTGTGACGCTGTGTGTCTGTGAGATTTCTGTCTCTGCCTGAGGG | 2014511 |
| QY | 2961 | ATGAGGGGTGCAGGGGTAGCTATCTTTGGAGATGGGTGC CAGGTGATCTGTTCGACATG | 3020 |
| Db | 201450 | ATGAGGGGTGCAGGGGTAGCTATCTTTGGAGATGGGTGC CAGGTGATCTGTTCGACATG | 2013911 |

| | | | |
|----|--------|---|--------|
| QY | 3021 | TGTCGCTGTGTGCACAAAGATATGTGTGCAGTCTGAAACATCTGTGCACACACGGCACTCTGT | 3080 |
| Db | 201390 | TGTGCTGTGTGCACAAAGATATGTGTGCAGTCTGAAACATCTGTGCACACACGGCACTCTGT | 201331 |
| QY | 3081 | GGGTGGCACTGAGACACTGTGGATAGAGGTGTGCATCCCGCTAAGCTCCCGGAGCGT | 3140 |
| Db | 201330 | GGGTGGCACTGAGACACTGTGGATAGAGGTGTGCATCCCGCTAAGCTCCCGGAGCGT | 201271 |
| QY | 3141 | GTGTACTGTGAGACAGAGCTGTATGTATGACTGACCTGTGTGAGGCAACATGGCGCTGTCT | 3200 |
| Db | 201270 | GTGTACTGTGAGACAGAGCTGTATGTATGACTGACCTGTGTGAGGCAACATGGCGCTGTCT | 201211 |
| QY | 3201 | GGAGAACTCCGTCGCTGTGCTGTACTGCTGTGTGTGGCGCTGGTTCTTTGGGGTAAT | 3260 |
| Db | 201210 | GGAGAACTCCGTCGCTGTGCTGTACTGCTGTGTGTGGCGCTGGTTCTTTGGGGTAAT | 201151 |
| QY | 3261 | TGCTAATATGATGTGTGTGCACAGGGCATATGACGAAGGTATGAACACGAGCCGGGCGCGTG | 3320 |
| Db | 201150 | TGCTAATATGATGTGTGTGCACAGGGCATATGACGAAGGTATGAACACGAGCCGGGCGCGTG | 201091 |
| QY | 3321 | GCTCACGCTGTATATCCAGCCCTTTTGGAGGGCCGAGCGGCGGATCACCTGAGGTCCG | 3380 |
| Db | 201090 | GCTCACGCTGTATATCCAGCCCTTTTGGAGGGCCGAGCGGCGGATCACCTGAGGTCCG | 201031 |
| QY | 3381 | GAGATCGAGGCCAGCTTGATCCAACTGAGAAACCCCGTCTTATCTATAAATATCAAAAT | 3440 |
| Db | 201030 | GAGATCGAGGCCAGCTTGATCCAACTGAGAAACCCCGTCTTATCTATAAATATCAAAAT | 200971 |
| QY | 3441 | TAGCTGTGTGTGTGGCGCGTGCCTGTATATCCAGCTATCCGGGAGACTGGGGCAGAAAA | 3500 |
| Db | 200970 | TAGCTGTGTGTGTGGCGCGTGCCTGTATATCCAGCTATCCGGGAGACTGGGGCAGAAAA | 200911 |
| QY | 3501 | TGCTGTGAAACCCGGAGGTGTGAGGTTCGGTGGAGGCCGAGATCGCGCATTTGCATCCAGC | 3560 |
| Db | 200910 | TGCTGTGAAACCCGGAGGTGTGAGGTTCGGTGGAGGCCGAGATCGCGCATTTGCATCCAGC | 200851 |
| QY | 3561 | CTGGGCAACAAAGGCAAACTCCGCTCTGAAAGAAAAAAGAAAAAAGGGTAAAGAA | 3620 |
| Db | 200850 | CTGGGCAACAAAGGCAAACTCCGCTCTGAAAGAAAAAAGAAAAAAGGGTAAAGAA | 200791 |
| QY | 3621 | CCAGTGAATGGGCAAGGGAGACTGATGATGAGAGTGGGCGATGCACTGTATGTCGTATGATC | 3680 |
| Db | 200790 | CCAGTGAATGGGCAAGGGAGACTGATGATGAGAGTGGGCGATGCACTGTATGTCGTATGATC | 200731 |
| QY | 3681 | TGTGTGTGAGAGAGAGATTGACAGATTGAGAGGCACTGTTTCACTGTGAATTTGAG | 3740 |
| Db | 200730 | TGTGTGTGAGAGAGAGATTGACAGATTGAGAGGCACTGTTTCACTGTGAATTTGAG | 200671 |
| QY | 3741 | AAACTTGAAGCTGTCTCTCCCTCCATGTGTGCCCCCTTAAAGTGAAGCCCTTCTCTGCT | 3800 |
| Db | 200670 | AAACTTGAAGCTGTCTCTCCCTCCATGTGTGCCCCCTTAAAGTGAAGCCCTTCTCTGCT | 200611 |
| QY | 3801 | CCTGTATTCCGAAACCTCTAGCTCCGCCANTGABACTGACCCCACTCCCTTCTCTGAACA | 3860 |
| Db | 200610 | CCTGTATTCCGAAACCTCTAGCTCCGCCANTGABACTGACCCCACTCCCTTCTCTGAACA | 200551 |
| QY | 3861 | CGCCCTTAGGCGACACTATGATGAGCCCGCCTTAAGGCCACACCCTTTGGGGCAGGCTC | 3920 |
| Db | 200550 | CGCCCTTAGGCGACACTATGATGAGCCCGCCTTAAGGCCACACCCTTTGGGGCAGGCTC | 200491 |
| QY | 3921 | CACCCCTATTTCTGTGGTACTCTTGTAAACCCCTTCAAAGTCAAGAGC-TTTTTTTTTT | 3979 |
| Db | 200490 | CACCCCTATTTCTGTGGTACTCTTGTAAACCCCTTCAAAGTCAAGAGC-TTTTTTTTTT | 200431 |
| QY | 3980 | TTTTTTTTTGGAGACAGTCTGTGCTCTCTCCAGGCTGAGTGCAGTGGCGTGTATCTCG | 4039 |
| Db | 200430 | TTTTTTTTTGGAGACAGTCTGTGCTCTCTCCAGGCTGAGTGCAGTGGCGTGTATCTCG | 200371 |
| QY | 4040 | CTCACTGCAACCTCTGACCTCCAGGTTCAAGTAAATTCTGATCTCTCACTCCTGAGTAG | 4099 |
| Db | 200370 | CTCACTGCAACCTCTGACCTCCAGGTTCAAGTAAATTCTGATCTCTCACTCCTGAGTAG | 200311 |
| QY | 4100 | CTGGAGTTTACAGTGTGGCGCAACACGCTGTGCTAATTTTGTCTTTAGTAGACAG | 4159 |


```

Db      200310 CTGGAGTTAAGAGGCGCGGCAACACGCTGGCTATTTTGTGTCTTTATAGACAG 200251
Qy      4160 GGTTCACCTTTGGCCAGGCTGTCTCAAACTCCCACTCAGTGATCCGCCACT 4219
Db      200250 GGTTCACCTTTGGCCAGGCTGTCTCAAACTCCCACTCAGTGATCCGCCACT 200191
Qy      4220 CGGCTCCAGAGTGTGGGGTTACAGCGGAGCACCGGCCAGCCCAAGTCAGAG 4279
Db      200190 CGGCTCCAGAGTGTGGGGTTACAGCGGAGCACCGGCCAGCCCAAGTCAGAG 200131
Qy      4280 CTCTTTAAGAGACTCTAACATGTAACTCTGACCTGGCCCTAATAAGTCAATTC 4339
Db      200130 CTCTTTAAGAGACTCTAACATGTAACTCTGACCTGGCCCTAATAAGTCAATTC 200071
Qy      4340 ACCCTCTCTGCTTCAGGCTTCAAGCCCACTCACTGAGGCTTGAAGCCCACTTTTGA 4399
Db      200070 ACCCTCTCTGCTTCAGGCTTCAAGCCCACTCACTGAGGCTTGAAGCCCACTTTTGA 200011
Qy      4400 CAGTTCAATCCCTAAAGCCCTGCTCCCTCCATCCCAAGGCTCAAGCCCAAGCTT 4459
Db      200010 CAGTTCAATCCCTAAAGCCCTGCTCCCTCCATCCCAAGGCTCAAGCCCAAGCTT 199951
Qy      4460 TGGCACTACCCCTGAGCTTGTCCAGAACTCTGTAACCAATTTTACCCTCATGTAGT 4519
Db      199950 TGGCACTACCCCTGAGCTTGTCCAGAACTCTGTAACCAATTTTACCCTCATGTAGT 199891
Qy      4520 CTACCAATTCAGAAATCTGTGAGGTCAGTAAAGTCACTGTAACCTTACCTGAGCTG 4579
Db      199890 CTACCAATTCAGAAATCTGTGAGGTCAGTAAAGTCACTGTAACCTTACCTGAGCTG 199831
Qy      4580 GGGCTCTCTGCTTGAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 4639
Db      199830 GGGCTCTCTGCTTGAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 199771
Qy      4640 GCCCTCTGCTTGAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 4699
Db      199770 GCCCTCTGCTTGAAGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTT 199711
Qy      4700 GAAACCAACCCCAAGCCCAAGCCCTGAGCCCGGCCCAACCAATTTTCCGTTCC 4759
Db      199710 GAAACCAACCCCAAGCCCAAGCCCTGAGCCCGGCCCAACCAATTTTCCGTTCC 199651
Qy      4760 AGAGCATGTTCTGGCAACATGATGTTCTGTGACCAACCCCTTAAACCCGCTC 4819
Db      199650 AGAGCATGTTCTGGCAACATGATGTTCTGTGACCAACCCCTTAAACCCGCTC 199591
Qy      4820 TGGAGAGCAACGAGACCTGGAGCTGGGGCGGGGAGAGACGCGGTCGATGACAGAG 4879
Db      199590 TGGAGAGCAACGAGACCTGGAGCTGGGGCGGGGAGAGACGCGGTCGATGACAGAG 199531
Qy      4880 CAGCCGATCATCATGATGATCGATCGATATGACACCCAGCCGTTGGAGCGCGGCT 4939
Db      199530 CAGCCGATCATCATGATGATCGATCGATATGACACCCAGCCGTTGGAGCGCGGCT 199471
Qy      4940 GTTGCTAAGGCGCAACAGCTCTGCTGCGGGCGGCTGTGTGATGATCAACAGTGGCT 4999
Db      199470 GTTGCTAAGGCGCAACAGCTCTGCTGCGGGCGGCTGTGTGATGATCAACAGTGGCT 199411
Qy      5000 C 5000
Db      199410 C 199410

```

```

RESULT 8
AC130782/c 200792 bp DNA linear HTG 22-NOV-2002
LOCUS Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
DEFINITION ordered pieces.
AC130782
AC130782.2 GI:25167101
VERSION HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE

```

```

ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 200792)
AUTHORS Banjamine, B., Blakeley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S., Carls, K., Coleman, B., Engle, J., Grant, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E., Laric, P., Lee, L.-S., O., Legaspi, R., Maduro, O. L., Maduro, V. B., Margulies, E. H., Mastello, C., Maekel, B., McDowell, J., Pennington, C., Pearson, R., Portnoy, M. E., Prasad, A., Reddix-Dugue, N., Schander, K., Schuler, M. G., Sison, C., Stanciu, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L., Weathers, K. D., Wiggins, L., Young, A. and Green, E. D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 200792)
AUTHORS Green, E. D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
REFERENCE 3 (bases 1 to 200792)
AUTHORS Green, E. D.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
COMMENT On Nov 22, 2002 this sequence version replaced gi:22218452.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nigr1.nih.gov
----- Project Information
Center project name: dmz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 19656 bases at least Q40
Consensus quality: 19783 bases at least Q30
Consensus quality: 19879 bases at least Q20
Insert size: 21500; agarose-fp
Insert size: 19892; sum-of-contigs
Quality coverage: 9.73x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 28306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length

```


| | | | |
|----|-------|---|-------|
| Db | 77243 | GAAgTCTGCGGCTgAGCTGGAGGAAATrCCCCACCCCTTAACTGGGGGACAGGGCAAT | 7718 |
| Oy | 1935 | GAGACCTGtGtAgGgGtGgCTCAGCAGCAGGAGAGAGAGtGtCTGtGgTCTTgCACc | 1994 |
| Db | 77183 | GAGACCTGtGtAgGgGtGgCTCAGCAGCAGGAGAGAGAGtGtCTGtGgTCTTgCACc | 77124 |
| Oy | 1995 | CACATCTTTCTCTGTCCCCCTCTTgCCCTGTCTGAGAGCTGCTAGACTCTATCTTCTGA | 2054 |
| Db | 77123 | CACATCTTTCTCTGTCCCCCTCTTgCCCTGTCTGAGAGCTGCTAGACTCTATCTTCTGA | 77064 |
| Oy | 2055 | ATTCTATAGtGcCGGgGtCTCAGGCGAGtGCGAGTGGtGgCCGtGCTTgTGGTCTCTG | 2114 |
| Db | 77063 | ATTCTATAGtGcCGGgGtCTCAGGCGAGtGCGAGTGGtGgCCGtGCTTgTGGTCTCTG | 77004 |
| Oy | 2115 | CTACCTGGGgGAAATAAgTAgGgGAGGAGGAGGAGtGgGTTAAgGgCTCCCGAGTCGc | 2174 |
| Db | 77003 | CTACCTGGGgGAAATAAgTAgGgGAGGAGGAGGAGtGgGTTAAgGgCTCCCGAGTgGc | 76944 |
| Oy | 2175 | CTGGGcCTCCCAACCTCTGTGACATTTCCCAATCCAGGTGACGCGGCATgGCTACAGCAG | 2234 |
| Db | 76943 | CTGGGcCTCCCAACCTCTGTGACATTTCCCAATCCAGGTGACGCGGCATgGCTACAGCAG | 76884 |
| Oy | 2235 | ACCCcCTGAGtGtGgGtGcCTGTGtGCTGAGTCAACAGcCTTgCTTgGGGgTCTACAG | 2294 |
| Db | 76883 | ACCCcCTGAGtGtGgGtGcCTGTGtGCTGAGTCAACAGcCTTgCTTgGGGgTCTACAG | 76824 |
| Oy | 2295 | TAAcCAGAACTCTGgGgTGGGAgGgTGTgGgGATTTGGAGAGACTGTCTCTGCGGCATAg | 2354 |
| Db | 76823 | TAAcCAGAACTCTGgGgTGGGAgGgTGTgGgGATTTGGAGAGACTGTCTCTGCGGCATAg | 76764 |
| Oy | 2355 | AGGcCTGTcCCCTGGGgAACTGTGTAGcCTTGGGCATGACTCCGgGACCGgGTGAATGT | 2414 |
| Db | 76763 | AGGcCTGTcCCCTGGGgAACTGTGTAGcCTTGGGCATGACTCCGgGACCGgGTGAATGT | 76704 |
| Oy | 2415 | GAGTCTGTCTGTGACTCTTGTGTGTGTGATGTGAATGTGGcCTGTGACTGTGCATGTGTGT | 2474 |
| Db | 76703 | GAGTCTGTCTGTGACTCTTGTGTGTGTGATGTGAATGTGGcCTGTGACTGTGCATGTGTGT | 76644 |
| Oy | 2475 | GTGTCTGGGgAGGgGgATGTcCTTTTCCATATCAAGGTGACTGTGCGCAGGTGCACTGAC | 2534 |
| Db | 76643 | GTGTCTGGGgAGGgGgATGTcCTTTTCCATATCAAGGTGACTGTGCGCAGGTGCACTGAC | 76584 |
| Oy | 2535 | CTTTTGAAGcCTGT | 2594 |
| Db | 76583 | CTTTTGAAGcCTGT | 76524 |
| Oy | 2595 | CACAGcTGT | 2654 |
| Db | 76523 | CACAGcTGT | 76464 |
| Oy | 2655 | GTGTGTGACTGGCATTTGATATATGACTGACAGTATCTGCACTTCTGTCCCTGAGGTCCGg | 2714 |
| Db | 76463 | GTGTGTGACTGGCATTTGATATATGACTGACAGTATCTGCACTTCTGTCCCTGAGGTCCGg | 76404 |
| Oy | 2715 | GGATTGGGTGCACAAAGTGTCTATCACCATGGAAAGCTGTGACTGTGTGTGTGTGTGTGA | 2774 |
| Db | 76403 | GGATTGGGTGCACAAAGTGTCTATCACCATGGAAAGCTGTGACTGTGTGTGTGTGTGTGA | 76344 |
| Oy | 2775 | GGCGATTATGTGATTGT | 2834 |
| Db | 76343 | GGCGATTATGTGATTGT | 76284 |
| Oy | 2835 | ACTACCTGAAGcCTGT | 2894 |
| Db | 76283 | ACTACCTGAAGcCTGT | 76224 |
| Oy | 2895 | TGTAAATGTACTGTATGT | 2954 |
| Db | 76223 | TGTAAATGTACTGTATGT | 76164 |
| Oy | 2955 | GGAGGgATAGAGGgTGCAGGgGTATGCTATCTGTGGAGATGGGTGCCAGGTACTGACTT | 3014 |
| Db | 76165 | GGAGGgATAGAGGgTGCAGGgGTATGCTATCTGTGGAGATGGGTGCCAGGTACTGACTT | 76106 |

[illegible]

| | | | |
|------------|---|---|----------------------------|
| OY | 4095 | AGTAGCTGGAGTTTACAGTGGGGGCGCCACACAGCGCTGGCTTAACTTTTGTGCTTAACTAGA | 4154 |
| Db | 75028 | AGTAGCTGGAGTTTACAGTGGGGGCGCCACACACACTGGCTAACTTTTGTGTTTAACTAGA | 74968 |
| OY | 4155 | GACAGGGTTTCACTTGTGGCCAGCGCTGGTCTCAACTCCCAACTCAGGTATCGCGC | 4214 |
| Db | 74968 | GACAGGGTTTCACTTGTGGCCAGCGCTGGTCTTCAACTCCCAACTCAGGTATCGCGC | 74909 |
| OY | 4215 | CACCTCGGCTCCCAAGTGTGGGGTTACAGGCTGAGGCCACCGCCCCAGCCCAAGT | 4274 |
| Db | 74908 | CACCTCGGCTCCCAAGTGTGGGGTTACAGGCTGAGGCCACCTGCCCCAGCCCAAGT | 74849 |
| OY | 4275 | CAGAGCTCTTATATGAGAGCTCTAACATGTAACTCTGACCCCTGACTTAACTAAGCAAT | 4334 |
| Db | 74848 | CAGAGCTC--TATAGGAGACTTACACTGTAACTCTGAGCTTAACTTAACTAAGCAAT | 74791 |
| OY | 4335 | TCCAAACCCCTTCTGCTCCAGCGCTGACCCCACTCACTGAGGCTGACCCCACTTCTT | 4394 |
| Db | 74790 | TCCAAACCCCTTCTGCTCCAGCGCTGACCCCACTCACTGAGGCTGACCCCACTTCTT | 74733 |
| OY | 4395 | GAGACCACTTTCATCTCTAAAGCCCTGGTCTCCCTCCATCTCCCAAGCTCCAGCCCCAC | 4454 |
| Db | 74730 | GAGACCACTTTCATCTCTAAAGCCCTGGTCTCCCTCCATCTCCCAAGCTCCAGCCCCAC | 74673 |
| OY | 4455 | AGCTTTGGACCTACCCCTGAGCTGTGCAGAGAACTCGTATCCGAATTTTAAACCTGCATG | 4514 |
| Db | 74670 | AGCTTTGGACCTACCCCTGAGCTGTGCAGAGAACTCTGTACCAATTTTAACTCTCCAG | 74611 |
| OY | 4515 | TAGTTCTAGCCCAATTCAGAGAACTGTGAGGTCCAGTTAGAGTCCAGTAACCTTACCTGA | 4574 |
| Db | 74610 | TAGTTCTAGCCCAATTCAGAGAACTGTGAGGTCCAGTTAGAGTCCAGTAACCTTACCTGA | 74551 |
| OY | 4575 | GCTTGGGCTGTCTCTTGAAGCTTGAAGCTTGAGAGGTGCACCTTATTTCTCCAG | 4634 |
| Db | 74550 | GCTTGGGCTGTCTCTTGAAGCTTGAAGCTTGAGAGGTGCACCTTATTTCTCCAG | 74491 |
| OY | 4635 | GCCCTGCCCCGCCCCCTCAGCATGTGACACACCAACCTCTAGCTGTCTGGCCTTGG | 4694 |
| Db | 74490 | GCCCTGCCCCGCCCCCTCAGCATGTGACACACCAACCTCTAGCTGTGTGGCCTTGG | 74431 |
| OY | 4695 | AGTCTGAACCCACCCCGACGCCAAGCGCCGCTGTAGCGCCGCCCAACCCATTTCGG | 4754 |
| Db | 74430 | AGTCTGAACCCACCCCGACGCCAAGCGCCGCTGTAGCGCCGCCCAACCCATTTCGG | 74371 |
| OY | 4755 | TTCCCAAGACATGTTCTCGCCCAACATGATGTTCTGTGACCAACCCCTCTAACACCGTG | 4814 |
| Db | 74370 | TTCCCAAGACATGTTCTCGCCCAACATGATGTTCTGTGACCAACCCCTCTAACACCGTG | 74311 |
| OY | 4815 | CCCTCTGGAGACACACAGACCTTGGAGCTTGGGGCGGGGAAACGCGCGGTGGATGAC | 4874 |
| Db | 74310 | CCCTCTGGAGACACACAGACCTTGGAGCTTGGGGCGGGGAAACGCGCGGTGGAGCAC | 74251 |
| OY | 4875 | AGCAGCAGCGCATCATCATATGATTCGCACTGCGCATATGACACACCAAGCGTGGACGGCC | 4934 |
| Db | 74250 | AGCAGCAGCGCATCATCATATGATTCGCACTGCGCATATGACACACCAAGCGTGGACGGCT | 74191 |
| OY | 4935 | GCGCTGTTGCTAAGAGCCCAACCAAGCTTACTCTGCGGGCGGTGTGTTGTGATCCACAGTGG | 4994 |
| Db | 74190 | GCGCTGTTGCTAAGAGCCCAACCAAGCTTACTCTGCGGGCGGTGTGTTGTGATCCACAGTGG | 74133 |
| OY | 4995 | CTGCTC 5000 | |
| Db | 74130 | CTGCTC 74125 | |
| RESULT 9 | | | |
| AC130188/c | | | |
| LOCUS | AC130188 | 176647 bp | DNA linear HTG 14-NOV-2002 |
| DEFINITION | Paojo anubis clone RP41-421p3, WORKING DRAFT SEQUENCE, 12 ordered pieces. | | |
| ACCESSION | AC130188 | | |
| VERSION | AC130188.2 GI:24960890 | | |
| KEYWORDS | HTG; HTGS_PHASE2; HTGS_DRAFT. | | |

SOURCE
ORGANISM
Papio anubis (olive baboon)
Bakayocra, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE
AUTHORS
1 (bases 1 to 176647)
Ahter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooke,S.
Carriaga,K., Coleman,B., Engle,J., Grante,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-O., Legaspi,R., Madero,Q.L., Madero,V.B.,
Marigales,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Parguirian,C., Pearson,R., Potnyoy,M.E., Prasad,A.,
Reddi-Dugue,N., Schandler,K., Schuler,M.G., Sison,C.,
Starrirpoy,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,B.D.
NSC Comparative Sequencing Initiative
REFERENCE
TITLE
2 (bases 1 to 176647)
Unpublished
AUTHORS
Green,B.D.
REFERENCE
TITLE
Direct Submission
JOURNAL
Submitted (08-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Galthersburg, MD 20877, USA
3 (bases 1 to 176647)
Green,B.D.
REFERENCE
TITLE
Direct Submission
JOURNAL
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Galthersburg, MD 20877, USA
On Nov 14, 2002 this sequence version replaced gi:22138439.
COMMENT

Genome Center
Center: NIH Intramural Sequencing Center
Center code: NSC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nhgri.nih.gov

Project Information
Center project name: der
Center clone name: 421P03

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171127 bases at least Q40
Consensus quality: 173424 bases at least Q30
Consensus quality: 174803 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 175547; sum-of-contigs
Quality coverage: 8.76x in Q20 bases; agarose-fp
Quality coverage: 9.33x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 6526: contig of 6526 bp in length
* 6527 6626: gap of unknown length

| | | | |
|---|--------|--------|------------------------------|
| * | 6627 | 18812 | contig of 12186 bp in length |
| * | 18813 | 18812 | gap of unknown length |
| * | 18913 | 61780 | contig of 42668 bp in length |
| * | 61781 | 61880 | gap of unknown length |
| * | 18181 | 97068 | contig of 35188 bp in length |
| * | 97069 | 97168 | gap of unknown length |
| * | 97169 | 125306 | contig of 28038 bp in length |
| * | 125207 | 125306 | gap of unknown length |
| * | 125307 | 127021 | contig of 1715 bp in length |
| * | 127022 | 127121 | gap of unknown length |
| * | 127122 | 129387 | contig of 2266 bp in length |
| * | 129388 | 129487 | gap of unknown length |
| * | 129488 | 132302 | contig of 2815 bp in length |
| * | 132303 | 132402 | gap of unknown length |
| * | 132403 | 135735 | contig of 3333 bp in length |
| * | 135736 | 135835 | gap of unknown length |
| * | 135836 | 143558 | contig of 7723 bp in length |
| * | 143559 | 143658 | gap of unknown length |
| * | 143659 | 170182 | contig of 26524 bp in length |
| * | 170183 | 170282 | gap of unknown length |
| * | 170283 | 176647 | contig of 6365 bp in length |

```

/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-421P3"
/clone_id="RP41"
1..6526
/note="assembly_fragment
clone_end:SP6
vector_side:left"
6627..18812
/note="assembly_fragment"
18913..61780
/note="assembly_fragment"
61881..97068
/note="assembly_fragment"
97169..125206
/note="assembly_fragment"
116911..176647
/note="clone overlaps with GenBank Accession Number
AC130786 clone RP41-94H5 (center project name das)"
125307..127021
/note="assembly_fragment"
127122..129387
/note="assembly_fragment"
129488..132302
/note="assembly_fragment"
132403..135755
/note="assembly_fragment"
135836..143558
/note="assembly_fragment"
143659..170182
/note="assembly_fragment"
170283..176647
/note="assembly_fragment
clone_end:TV
vector_side:right"

```

ORIGIN

| | | | | |
|----------------------------|-------|---------------------|------------|----------------|
| Query Match | 20.7% | Score 1035.2; | DB 2; | Length 176647; |
| Best Local Similarly | 88.1% | Pred. No. 1.3e-260; | | |
| Matches 1256; Conservative | 0; | Mismatches 123; | Indels 47; | Gaps 10; |

| Qy | Db | Qy | Db |
|---|--|--|---|
| 658 | 1410 | 718 | 1350 |
| ATGATCACTCTTGACATATCTATATGACAAGACCTAAAGCATACAGACTTAACCT | ATCATCCCTTGATTATGTGTATGCTGTGACAGATACCTAAAGCATATAGACTCCACCT | TTAAATATGACAGTTTGCGGCACAGCACCGCTGCTCATGCTGTAAATTCACGACATTYGGGA | TTAAATATGACAGTTTGCGGCACAGCATGCTGCTCATGCTGTAAATTCACGACATTYGGGA |

| | | | |
|----|------|--|------|
| Qy | 778 | GGCAGAGTGGGTAAATCACTGAGGGCAGAGGTTTGAACACAGCTGGCCAAACATGGTG | 837 |
| Db | 1290 | GGCAGAGTGGGTAAATCACTGAGGGCAGAGGTTTGAACAGACACAGCTGGCCAAATGGTG | 1231 |
| Qy | 838 | AAATCTGTCTTTTAACTAAAAAATAAAAAATAAAAATGAGCCGGGTGTCTGGGGG | 897 |
| Db | 1230 | AAATCCCATCTCTACT-----AAAAAATAAAAAATAAAAATGAGCCGGGCATGTGGGG | 1176 |
| Qy | 898 | CACACCTGTAAATCCACGCTATGCTGAGGCTGAGGACGAGAGTCACTTGA--ACCTGG | 955 |
| Db | 1175 | --CACCTGTAAATCCCACTACTCCAGAGGCTGAGGCTGAGAGTCCGTTGAACACAGG | 1118 |
| Qy | 956 | AGGCGGAGGTTTGCAGTGGGCGAGATCAATCAACCCGCCCTCAGCCCTGGGCGAAGAGCA | 1015 |
| Db | 1117 | AGGAGAGGTTTGCAGTGGGCGAGATCAACACTGCCCTCCAGCCTGGGGTGAACAGCA | 1058 |
| Qy | 1016 | AGACTGTCTCAAAATTAATTAATAAACAAGACAGAGTTGTGTACTTAACTTA | 1075 |
| Db | 1057 | AGACTGTCTCTCAAAATTAATTAATAACA-----AACACAGATTGTATTACTTAAGTA | 1002 |
| Qy | 1076 | TATCT-----AAAAAATATGCTGTCAACAATAGAGCAGAAGTGAATTAAGAAAAATAA | 1132 |
| Db | 1001 | TACCTAAAAAAAATATGCTGTCAACAGATAGAGCAGAAGTGAATTAAGAAAAATAA | 942 |
| Qy | 1133 | ATGGGCCCAAGAACTCTAAGGTATTTTGAACAAATCATTCAGAACTTTAAAAAAGAAAGA | 1192 |
| Db | 941 | ACAGGCTGAGAACTCTAAGGTATATTGGAACAAATCATCTCAGAACCTTTAAAAAAGAGAGA | 882 |
| Qy | 1193 | ATCACAAGGCAATAGAAAGAC---AGGAGAGAAACAGGGAGACAGAAACACTGTGGCCC | 1248 |
| Db | 881 | ATCACAAGGCAATAGAGAGACAGAGACGAGAAATGGGGAGACAGCACACTTGTGGCCC | 822 |
| Qy | 1249 | AAGGAGAACAAAACAAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGTGTGAG | 1308 |
| Db | 821 | AAGGAGAGAGAAACAAGGCTCTTAAGACA-----GGAGAGAGAGAGAGAGAGAGAGAG | 767 |
| Qy | 1309 | AGACACACAGAGAAAAAGACAGAGACAGAGAGACAGACACAGAGAGACAGAGAGCCAGAA | 1368 |
| Db | 766 | CGCACTTGACAGAGAGACAGCCACAGACAGAGACAGAGAGACAGAGAGACAGAGAGGTAGA | 707 |
| Qy | 1369 | GGGATAGAAAAAGAGAC---GAGGGGTGAGAGAGACAGAGATTTTAGAGAGACTGAGA | 1425 |
| Db | 706 | GGGAGAGAGAAAGAGAGGGGATGAGAGAGACAGAGATTTTAGAGAGAGACTGAAA | 647 |
| Qy | 1426 | AAAGTACCCGAGGGAGAACACAGAGAGATGGAAGAAAGACTCTG----- | 1469 |
| Db | 646 | AAAGACAGCCCTGGGAGAACACAGAGAGATGGAAGAAAGACTCTGAAAAAAAAAAAAAGAAA | 587 |
| Qy | 1470 | -AGAAAAACCAAGACAAAGATGAAAAAGAGATATCGAGGGTGAACAGACAGTGGTG | 1528 |
| Db | 586 | AAAGAAAAAACAGAGCAAAAGATGAAAAAGAGATATCGAGCGGGAACAACAGTGGTG | 527 |
| Qy | 1529 | AATGAGCAAAATCCAGAGAGAAAGAAAGCAAGCAATCCAGGCCCAAGAAATAGTACCCAGAG | 1588 |
| Db | 526 | AATGAGCAAAACGAGAGAAAGAAAGCAAGCAATCCAGACACAGAAACAGTGAACAAGAG | 467 |
| Qy | 1589 | TTGGTGAAGAACCAAGATCTTTAAGGCTGGGGAGGACAGGAAAGGGGCTGGCTGGCTTC | 1648 |
| Db | 466 | TTGGTGAAGAACCAAGATCTCTTAAGGCTGGGGAGGACAGGGAAGGGGCTGGCTGGCTTC | 407 |
| Qy | 1649 | GGAGACCCCTCCCATTTCTCGGGCCAGGGAGGTAGGAGTGAACATTCGGAATTGGGTGG | 1708 |
| Db | 406 | GGAGACCCCTCCCATCTCTCGAGCCAGGGAGGTAGGGAATGTGATATTCGGGAACAGGTGG | 347 |
| Qy | 1709 | GGGTGTCTCTGGGGGTGAGATAGGGGGAGACAGAGAGAGCTATTGCTTAAGCCCGATAGG | 1768 |
| Db | 346 | GGGTGTCTCTGGGGGTGAGATAGGGGGAGACAGAGAGAGGATGTGCTTAAGGCCCAATAGG | 287 |
| Qy | 1769 | CACCTCAATTCGCCGGGAATGTGGCCCAAGGAGCAGTGGGTGTTATTAATCAGAGCCCGGT | 1828 |
| Db | 286 | CACCTCAATGGCCCGGGAAATGTGCCCCAGGGAGCAGTGGGTGTTATTAATCAGAGCCCGGT | 227 |
| Qy | 1829 | GCCCAAGGCCCAGAGAGAGCAGTGGCCAGAGAAAGCAACAGGCTTGAGAAAGTCTGCGGCTG | 1888 |

|||||
Db 226 GCCCAGAGCCGAGAGGCGGTGGCCAGAGAGTGCAGCCCTGAGAAATCCGTGGCTG 167
1889 AGCTGGAGAGCAATCCCGCCCTTACCTGGGGGAGACAGGCGAATGACCTGGTGAGG 1948
Db 166 AACTGGAAGGAATCCCGCCCTTACCTGGGGGAGACAGGCGAATGACCTGGTGAGG 107
Qy 1949 GTGGCTCAGCAGCAGGAGAGAGTGTCTGTGCGTCTGCACCCACATCTTTCTCTG 2008
Db 106 GTGGCTCAGCAGCAGGAGAGAGTGTCTGTGACCTCTGCACCCACATCTTTCTCTG 47
Qy 2009 TC--CCCTCCTTCCCTGTCTGTGAGCGTGTGACCTCTTATCTTCT 2052
Db 46 TCTTCCCTCCTCGCCCGTGTGAGGTTGCTAGACTCTTATCTTCT 1
RESULT 10
AC137591/c 32799 bp DNA linear PRI 05-DEC-2002
LOCUS Homo sapiens chromosome X clone XX-B6cos map Xp22-PAR, complete
DEFINITION sequence.
AC137591
VERSION AC137591.1 GI:25446705
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Rump,A., Rosenchal,A., Drescher,B. and Schatkevov,R.
TITLE 1 (bases 1 to 32799)
JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE
AUTHORS lagemann,D. and Platzter,M.
TITLE 2 (bases 1 to 32799)
JOURNAL Submitted (26-NOV-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE
AUTHORS lagemann,D. and Platzter,M.
TITLE 3 (bases 1 to 32799)
JOURNAL Submitted (05-DEC-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
COMMENT

Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de

Project Information
Center project name: B6
Center clone name: XX-540F24

Summary Statistics
Sequencing vector: pUC18, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 32765 bases at least Q40
Consensus quality: 32770 bases at least Q30
Consensus quality: 32770 bases at least Q20
Quality coverage:13.12x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source
Location/Qualifiers
1..32799
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp22-PAR"
/clone="XX-B6cos"
537..642
/note="single stranded/single chemistry region"
2807..2820
/note="single stranded/single chemistry region"
2913..2927
/note="low quality region"
3166..3173
/note="single stranded/single chemistry region"
5131..5234
/note="single stranded/single chemistry region"
5235..5317
/note="single clone coverage"
5266..5317
/note="low quality region"
5266..5317
/note="single clone coverage"
5366..5431
/note="single clone coverage"
5432..5518
/note="single stranded/single chemistry region"
6151..6395
/note="single stranded/single chemistry region"
6303
/note="single stranded/single chemistry region"
6382
/note="single stranded/single chemistry region"
7106..7137
/note="single stranded/single chemistry region"
9265..9528
/note="single stranded/single chemistry region"
9669
/note="T substituted in clone: XX-P99cos"
/replace="c"
10317
/note="deleted in clone: XX-540F24"
/replace=""
10897
/note="G substituted in clone: XX-1Bcos"
/replace="t"
11087..11090
/note="deleted in clone: XX-540F24"
/replace=""
11259
/note="T substituted in clone: XX-540F24"
/replace="c"
11744..11756
/note="low quality region"
11766..11880
/note="single stranded/single chemistry region"
11834
/note="T substituted in clone: XX-540F24"
/replace="c"
11892..11900
/note="deleted in clone: XX-1Bcos"
/replace=""
12197..12200
/note="deleted in clone: XX-1Bcos"
/replace=""
12256..12259
/note="deleted in clone: XX-1Bcos"
/replace=""
12277
/note="A substituted in clone: XX-540F24"
/replace="c"

| | |
|--------------|--|
| variation | 12290..12291 /note="deleted in clone: XX-540F24" /replace="" |
| variation | 12312 /note="G substituted in clone: XX-540F24" /replace="a" |
| variation | 13352 /note="G substituted in clone: XX-1BCo8" /replace="a" 13768 |
| variation | /note="T substituted in clone: XX-540F24" /replace="g" 14138 |
| variation | /note="T substituted in clone: XX-540F24" /replace="a" 14664 |
| variation | /note="C substituted in clone: XX-540F24" /replace="c" 15355 |
| variation | /note="deleted in clone: XX-1BCo8" /replace="" 15447..15448 |
| variation | /note="deleted in clone: XX-1BCo8" /replace="" 15481..15483 |
| variation | /note="deleted in clone: XX-540F24" /replace="" 15868 |
| variation | /note="A substituted in clone: XX-540F24" /replace="g" 15961 |
| variation | /note="C substituted in clone: XX-540F24" /replace="c" 16252 |
| variation | /note="C substituted in clone: XX-540F24" /replace="c" 16334 |
| variation | /note="A substituted in clone: XX-1BCo8" /replace="g" 16684 |
| variation | /note="C substituted in clone: XX-540F24" /replace="a" 16793 |
| variation | /note="G substituted in clone: XX-540F24" /replace="a" 16817 |
| variation | /note="C substituted in clone: XX-540F24" /replace="g" 17129 |
| variation | /note="C substituted in clone: XX-540F24" /replace="c" 17489 |
| misc_feature | /note="C substituted in clone: XX-540F24" /replace="g" 18712..18717 |
| misc_feature | /note="single stranded/single chemistry region" 21509..21686 |
| variation | /note="single stranded/single chemistry region" 21628 /note="A substituted in clone: XX-540F24" /replace="g" 21652 |
| variation | 21652 /note="G substituted in clone: XX-540F24" /replace="a" 22356..22558 |
| misc_feature | /note="single stranded/single chemistry region" 22529..22586 |
| misc_feature | /note="single clone coverage" 22606 |
| unre | /note="low quality region" 22606 |

| | | | | | | |
|----|-------------|---|---------------------|-----------------|---------------|---------|
| | Query Match | 9.2%; | Score 461.6; | DB 9; | Length 32799; | |
| | Similarity | 63.9%; | Pred. No. 6.3e-110; | | | |
| | Matches | 838; | Conservative 0; | Mismatches 444; | Indels 30; | Gaps 8; |
| Oy | 325 | GGGtgggtcATATCATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTAGACCGAG | 384 | | | |
| Dd | 6412 | GGTATTACTCTAATGCATATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGGGAG | 6353 | | | |
| Oy | 385 | TCTGCTCTGTCCGCCAGGTGAGTGCAGTGGCGGATCTCGCTCATTGCAAGCTCCG | 444 | | | |
| Dd | 6352 | TCTGCTCTGTCCGCCAGGTGAGTGCAGTGGCGGATCTCGGCTCATGCAAGCTCCA | 6293 | | | |
| Oy | 445 | CCTCCCGGGTTCAGGCCATTCTCTGCTCAGCTCCCAGTAGCTGGACTAGACGCC | 504 | | | |
| Dd | 6292 | CCTCCCGGGTTCAGGCCATTCTCTGCTCAGCTCCCAGTAGCTGGACTAGACGGTTC | 6233 | | | |
| Oy | 505 | CCGCACTACGCCCGGGTAATTTTTTTTGTATTTTGTAGAGACGGGGTTTCAACGGTTT | 564 | | | |
| Dd | 6232 | CCGCACTACGCCCGGGTAATTTTTTTTGTATTTTGTAGAGACGGGGTTTCAACGGTTC | 6173 | | | |
| Oy | 565 | AGCCGGATGGCTCGATCTCTGACCTCGTATCCGCCCGCTCGGCTCCCAAAGTGC | 624 | | | |
| Dd | 6172 | TTCCAGATGGTCTCGATCTCTGAGTCAATGATCCACCGETTCGGGCTCCAAAGTGC | 6113 | | | |
| Oy | 625 | TGGATTACAGGGCTGACCCGCCGCCCGCCATGATCATCTTTCTTGACTATGCTATG | 684 | | | |
| Dd | 6112 | TGGATTACAGGGCTGACCCCTATGTCAGGCTGACATCTCAATTTTAAACATTTGAG | 6053 | | | |
| Oy | 685 | TGAACAATCAATGAAGCATCAGACTCAOCCCTTAATAATGACGTTTGGGCCAGGCACC | 744 | | | |
| Dd | 6052 | ATTAAATTAAAGTGTGTGCCCGTGTAGA-----AAACTGGCACTCTGGCCAGGCACA | 5999 | | | |
| Oy | 745 | GTGCTCATGCTGTAAATTCAGACCTTTGGAGCGAGAGTGGGTGATCACTTGAAGC | 804 | | | |
| Dd | 5998 | GTGCTCATGCTGTCAATCTCAGACCTTTGGGAATTTGAGGTGGCGAATCAC--GAGGT | 5941 | | | |
| Oy | 805 | CAGAGTTTGAACCAAGCTGGCCACAATGTGAACTGTCTTTACTAAAAAAA | 864 | | | |
| Dd | 5940 | CAGGAATTCAGAACCAAGCTGGCCACAACGGTGAACCCCATCTCTACTA----- | 5891 | | | |
| Oy | 865 | AAAAAAAAAAATCAGCCGGGTGTGTGGGGCACCTGTATCCAGCTATGCTGA | 924 | | | |
| Dd | 5890 | --AATATACAAAATTAAGCTGGAGTGGTGGCGATGCTGTATCCAGCACTTGGGA | 5833 | | | |
| Oy | 925 | GGCTGAGGCAAGAGTCACTTGAACCTCGAGCGGAGGTTGCACTGGCCGAGATCAC | 984 | | | |
| Dd | 5832 | GGCTGAGGCAAGAGATCTGTTGAACCAAGAGGTGAGGTTGATGAGCCGAGATGCC | 5773 | | | |
| Oy | 985 | ATCACCGCCTCCAGCCTGGCGACAGACAGACTGTCTCAATTAATTAATA | 1044 | | | |
| Dd | 5772 | GCTCTGCACTCCAGCCT--GGTGACAGACAGACTCT-----TAAAAAGAAAGAAAG | 5720 | | | |
| Oy | 1045 | AACGACAAGCATTTTGTGTACTTATCTTAATAAAAAAAATGCTCAACAAT | 1104 | | | |
| Dd | 5719 | AAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAAAGAAAGAAAGAAAG | 5660 | | | |
| Oy | 1105 | AGAGCAGAGTGAATTAAGGAAATTAATGGGCCAAGAACTTAAGTATTTTGA | 1164 | | | |
| Dd | 5659 | AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG | 5601 | | | |
| Oy | 1165 | ATCATTCAGACTTTTAAAAAGAAATCAACAGAGCATAGAAAGACAGGAGAA | 1224 | | | |
| Dd | 5601 | GAAGAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAATAGAAAGAAAGAAAG | 5541 | | | |
| Oy | 1225 | GGGAGACAGAAACCTGTGTG--CCCAAGAGAACAAACAAAGCTCTTAAGACAGACAG | 1283 | | | |
| Dd | 5540 | AGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAAAGAAAGAAAGAAAGAAAG | 5481 | | | |
| Oy | 1284 | AGGAGAGAGAGAGAGTGTGTGAGACAGACAGAGAAAGAAAGACAGAGAGAGACA | 1343 | | | |
| Dd | 5480 | GGAGGAG | 5421 | | | |
| Oy | 1344 | GAGACAGAGACAGAGAGGCGAGAGGATAGAAAGAGA-GAGGGGTGAGAGACA | 1402 | | | |


```

Db      5420 TGGAAAGAAAGAAAGCAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5361
QY      1403 CCGAGATTTGAGAGAGACTCAGAAAGATAGCCAGAGAGAAACCAAGAGATGAGAAAG 1462
Db      5360 AAGAAAGAAATGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 5301
QY      1463 GACTCTGAGAAAAAACCAGAGCAAAAGATGAGAAAGAGTATCGAGGGTGAACACAGAC 1522
Db      5300 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5241
QY      1523 TGGTGAATGAGCAAAATGCAAGAAAGAAAGCAAGCAATCCAGCCCAAGAAATGATGAC 1582
Db      5240 AGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5181
QY      1583 CCAAGATTGTTGTAAGCCAGATCTTAAGCTGCTGGGGAGAGCGAGAAAGGG 1634
Db      5180 GGAAGGAAGGAAGAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAG 5129

```

```

RESULT 11
AC149102/c 190603 bp DNA linear HTG 13-MAY-2004
LOCUS Papio anubis clone RP41-113H15, WORKING DRAFT SEQUENCE, 17
DEFINITION
unordered pieces.
AC149102
AC149102.1 GI:47155599
VERSION HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS Papio anubis (Olive Baboon)
SOURCE Papio anubis
ORGANISM

```

```

REFERENCE
AUTHORS Antonellis,A., Ayele,K., Benjamin,B., Blakeley,R.W.,
Boutfair,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hulse,B.,
Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margillies,E.H., Mastiello,C., Maekert,B., McDowell,U.,
Mullikin,J.C., Paguitigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddy-Dugue,N., Schandler,K., Schaefer,M.G., Shah,K., Sison,C.,
Stantirpop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
MISC Comparative Sequencing Initiative

```

```

1 (bases 1 to 190603)
2 (bases 1 to 190603)
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

```

```

Submitted (13-MAY-2004) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

```

```

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoemhgrl.nih.gov
Project Information
Center project name: gzf
Center clone name: 113H15

```

```

Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 178234 bases at least Q40
Consensus quality: 180810 bases at least Q30
Consensus quality: 182566 bases at least Q20
Insert size: 223000; agarose-fp
Insert size: 189003; sum-of-contigs
Quality coverage: 10.1ix in Q20 bases; agarose-fp
Quality coverage: 11.9ix in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently

```

* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2285: contig of 2285 bp in length
2286: gap of unknown length
2385: gap of 3769 bp in length
2386: gap of 6154 bp in length
6155: gap of unknown length
6254: gap of unknown length
6255: contig of 4391 bp in length
10645: gap of unknown length
10646: gap of unknown length
10745: gap of 3853 bp in length
14608: gap of unknown length
14609: gap of 6769 bp in length
14709: gap of 6769 bp in length
21477: gap of unknown length
21577: gap of 8249 bp in length
21578: gap of 8249 bp in length
29827: gap of unknown length
29926: gap of unknown length
34784: gap of 4858 bp in length
34785: gap of unknown length
34885: gap of 5555 bp in length
40439: gap of unknown length
40440: gap of unknown length
40539: gap of 7356 bp in length
40540: gap of 7356 bp in length
47895: gap of 7356 bp in length
47995: gap of unknown length
47996: gap of 15311 bp in length
63306: gap of unknown length
63406: gap of 17150 bp in length
63407: gap of unknown length
80556: contig of 17150 bp in length
80557: gap of unknown length
80558: gap of 8113 bp in length
80559: gap of 8113 bp in length
88769: gap of unknown length
88870: gap of 8993 bp in length
97862: gap of 8993 bp in length
97863: gap of unknown length
97962: gap of 11438 bp in length
97963: gap of 11438 bp in length
109400: contig of 11438 bp in length
109401: gap of unknown length
109500: gap of unknown length
133884: contig of 24384 bp in length
133885: gap of unknown length
133985: gap of unknown length
166228: contig of 32644 bp in length
166629: gap of unknown length
166729: contig of 23875 bp in length.

```

FEATURES

source

```

1..190603
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-113H15"
/clone_lib="RP41"
/note="BAC resource: http://bacpac.choxi.org/"
1..2285
/note="assembly_fragment"
2386..6154
/note="assembly_fragment"
6255..10645
/note="assembly_fragment"
10746..14608
/note="assembly_fragment"
14709..21477
/note="assembly_fragment"
21578..29826
/note="assembly_fragment"
29927..34784
/note="assembly_fragment"
34885..40439
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
40540..47895
/note="assembly_fragment"
47996..63306
/note="assembly_fragment"
63407..80556
/note="assembly_fragment"
80657..88769

```

```
misc_feature      /note="assembly_fragment"  
                  88870..97862  
                  /note="assembly_fragment  
clone_end:17  
vector_side:left"  
misc_feature      97963..109400  
                  /note="assembly_fragment"  
misc_feature      109501..133884  
                  /note="assembly_fragment"  
misc_feature      133985..166628  
                  /note="assembly_fragment"  
misc_feature      166729..190603  
                  /note="assembly_fragment"
```

ORIGIN

| | | |
|----|---|---|
| | Query Match | 9.1%; Score 457.2; DB 2; Length 190603; |
| | Best Local Similarity | 77.8%; Pred. No. 1,je-108; |
| | Matches | 582; Conservative 0; Mismatches 153; Indels 13; Gaps 2; |
| QY | 346 TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGGAGCTCTCGCTCTGTCCGCAAGGCT | 405 |
| Dy | 43056 TTGTGTTTCCTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGCTCTCGCTCTGTCCGCAAGGCT | 42997 |
| QY | 406 GGAGTGCAGTGGCGGGATCTCGGCTCCTGCAAGCTTCGGCTCCCGGATTACGCCATTTC | 465 |
| Dy | 42996 AGAATGCAGTGGCGGGATCTCAGCTCATCAAGCTCCGGCTCCCGGATTACGCCATTTC | 42937 |
| QY | 466 TCCGAGCTCAAGCCTCCCAAGTAGTGGAGCTACAGAGGCGCCGGCACTAGCGCCGGCTAAAT | 525 |
| Dy | 42936 TCTCGGCTCAAGCTCCCGAGTAGTGGAGCTACAGAGGCGCCGGCACTAGCGCCGGCTAAAT | 42877 |
| QY | 526 TTTT---GTAATTTTGTAGAGACGGGGTTTACCGTITTTTACCGGGATGAGCTCGATC | 583 |
| Dy | 42876 TTTT---GTAATTTTGTAGAGACGGGGTTTTCGGCTGTTTACGAGGGTGTCTGATC | 42811 |
| QY | 584 TCTGACCTGTGATCCGCCCGCTCGGCTCCCAAAGTCTGGGATTACAGGCGTGAAC | 643 |
| Dy | 42816 TCTGACCTGTGATCCGCCCGCTCGGCTCCCAAAGTCTGGGATTACAGGCGTGAAC | 42757 |
| QY | 644 CACGGGCGCGGGCATGATCATCTCTTGACTATGCGATGTAGACAAGTACCCTAAGCCA | 703 |
| Dy | 42756 CAGGCGCGCGGGCATGATCATCTCTTGACTATGCGATGTAGACAAGTACCCTAAGCTA | 42697 |
| QY | 704 TCAGACTCTACCTTTTAAATATGACATGTTGGGCGAGCACCTGTGCTCATGCTGTAAAT | 763 |
| Dy | 42696 CAACCTGAAATTCATTTAAAGTGTGTATAGGCGAAGTGGGTGGCCATGCTGTGATC | 42637 |
| QY | 764 CCAGCACTTTGGAGGCGAGAGTGGTGAATCACTTGAGGCCAGAGTTTGAGACCAACC | 823 |
| Dy | 42636 CCAGCACTTTGGAGGCGAGAGTGGTGGAGCACTTGAGGTCAGAGTTGAGAACCAACC | 42577 |
| QY | 824 TGCGCAACATGGTGAACCTGTGCTTTTACATAAAAAAAAAAAAAAAAAAAAAATTCAGC | 883 |
| Dy | 42576 TGCTCAATGTGTGAACCCCGTCTTACT-----AAAAAAAAACAAATTCAGC | 42528 |
| QY | 884 CGGATGTCTGTGGGCAACCTGTAAATCCAGCTATGTGTGAGGCTGAGGCAAGAGTCA | 943 |
| Dy | 42527 TGGGCGTGTGTGGGCAACCTGTAAATCCAGCTATGTGTGAGGCTGAGGCAAGAGTCA | 42466 |
| QY | 944 CTTGAACCTTGGAGGCGAGGTTTGACAGTGGGCGAGATCATCATCCGCCCTTCAGCCTG | 1003 |
| Dy | 42467 CTTGAACCTTGGAGGCGAGGTTTGCAATGAGCAAGAATTTGGCAACCCCATCTCAACCTG | 42408 |
| QY | 1004 GGGCAGAGAGCAACCTGTCTCAATTAATTAATTAACAAAGCAACAGCAGTTGTT | 1063 |
| Dy | 42407 GGGCGCAGAGCAACCTGTCTCAAAAAAAAAATTAATTAATTAAGTGTCTCATGATTTGAA | 42348 |
| QY | 1064 GTACCTTAGTTATATCTTAAAAAAAAT 1091 | |
| Dy | 42347 GGACTAAATATGACATCTTAAAAAAT 42320 | |

| | | | | |
|------------|---|-------------|--------|-----------------|
| AC109997 | 141923 bp | DNA | linear | HTG 09-FEB-2002 |
| LOCUS | | | | |
| DEFINITION | Homo sapiens chromosome 5 clone CTD-207601, WORKING DRAFT SEQUENCE, | | | |
| | 3 unnumbered pieces. | | | |
| ACCESSION | AC109997 | | | |
| VERSION | AC109997.1 | GI:18642721 | | |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN. | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| REFERENCE | 1 (bases 1 to 141923) | | | |
| AUTHORS | DOE Joint Genome Institute. | | | |
| TITLE | Sequencing of Human Chromosome 5 | | | |
| JOURNAL | Unpublished | | | |
| REFERENCE | 2 (bases 1 to 141923) | | | |
| AUTHORS | DOE Joint Genome Institute. | | | |
| TITLE | Direct Submission | | | |
| JOURNAL | Submitted (09-FEB-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94558, USA | | | |
| COMMENT | -----Genome Center | | | |

COMMENT

```

Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 657070
Center clone name: CITB-HI_207601
-----
Summary Statistics
Consensus quality: 139579 bases at least Q40
Consensus quality: 140059 bases at least Q30
Consensus quality: 140264 bases at least Q20
Estimated insert size: 140000; agarose-1f estimation
Estimated insert size: 141733; sum-of-contigs estimation
Quality coverage: 8.57 in Q20 bases; agarose-1f estimation
Quality coverage: 8.47 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
    *   1      1097: contig of 1097 bp in length
    *       1
    *   1098      1197: gap of unknown length
    *       *
    *   1198      18670: contig of 17473 bp in length
    *       *
    *   18671      18770: gap of unknown length
    *       *
    *   18771      141923: contig of 123153 bp in length.
FEATURES
Source
    1..141923
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="5"
        /clone="CTD-207601"
        /clone_lib="Caltech human BAC library D"
ORIGIN
Query Match          9.1%; Score 456.6; DB 2; Length 141923;
Best Local Similarity 79.0%; Pred. No. 1,6e+108;
Matches 573; Conservative 0; Mismatches 139; Indels 13; Gaps 2;
OY      326 GGCGGCTCTATGATCATCTCTTTTGTGGAGCGGAGT 385
           |||||
Db      91006 GATGGTCATTCTCCTCAGGTTTTTTTGTTTGTGGAGCGAGT 91065
OY      386 CTGGCTCTGCGCCAGGCTGAGTGCAGTGGGGGAGTCTGGGCTACTGCAGCTCCGC 445
           |||||
Db      91066 CTTGCTCTTGCGCCAGGCTGAGTGCAGTGGGGGAGTCTGGGCTACTGCAGCTCCGC 91125
OY      446 CTCGCCGGTTCAGCGCAATTCCTCGCTCAGCTCCCAAGTAGCTGGAGCTACAGCGCC 505

```

| Db | 91126 | CTCCAGATTACAGCCATTCTCTGCTCAGCTCCCGAGTACTGGAGCTACAGGCGCC | 91185 |
|------------|--|--|----------------------------|
| OY | 506 | CGCCACTACGCCCGGCTAAATTTTTTTGTAATTTTTTAAGTAGAGACGGGGTTTCAACCGTTTTA | 565 |
| Db | 91186 | CGCCACACAGCGCCGGCTAAATTTTTTTGTAATTTTTTAAGTAGAGACGGGGTTTCAACCGTTTTA | 91245 |
| OY | 566 | GCCGGGATATGGCTTGATCTCTCTGACCTCTGATGATCCGCGCGCTCTGGCTCTCCCAATGTCT | 625 |
| Db | 91246 | GCCAGGATATGGCTTGATCTCTCTGACCTCTGATGATCCGCGCGCTCTGGCTCTCCCAATGTCT | 91305 |
| OY | 626 | GGGATTAACAGCGCTGAGACCAACCGGCGCCGACATCTTCTTGAATCTGATGATGATG | 685 |
| Db | 91306 | GGGATTAACAGCGCTGAGACCAACCGGCGCCGACATCTTCTTGAATCTGATGATGATG | 91364 |
| OY | 686 | GACAAATACCTAAAGCCATGACATCTTACCCCTTAAATATGACATTTGGGCCAGGACCG | 745 |
| Db | 91365 | GCTCATCTAAATGAAGAAATTTATGACAGGGGTTTAAATACTAGGACGGGATGGCATGA | 91424 |
| OY | 746 | TGCGTCAATGCTGTAATTTCCAGACCTTTGGGAGCAGAGGTGGGTGATCATCTTAGAGCC | 805 |
| Db | 91425 | TGCGTCAATGCTGTAATTTCCAGACCTTTGGGAGCAGAGGTGGGTGATCATCTTAGAGCC | 91484 |
| OY | 806 | AGGATTTAGAGACCAAGCTGGCCCAACATGATGAAACCTGTCTTTCCTAAATAAAAA | 865 |
| Db | 91485 | CGAGTTTCAGACCAAGCTGGCCCAACATGATGAAACCTGTCTTTCCTAAATAAAAA | 91532 |
| OY | 866 | AAAAAATAAAAATACAGCCCGGGTGTCTGTGGGGACACCTGTAAATCCAGCTATCTGAG | 925 |
| Db | 91533 | AAAAATCAAAAATTTAGCAGAGCTTAAATGGTGTGTCTGTATGCCAGCTACTCGGAG | 91592 |
| OY | 926 | GCTGAGGACGAGAGTCACTTTGAAACCTTGGAGCGAGGTTGCAATGGGCCAGATCA | 985 |
| Db | 91593 | TCTAGGCGATGAGATCACTTGAACCTGGAGGTTGCAATGGGCCAGATTA | 91652 |
| OY | 986 | TCATCGCCCTCCAGCTGGGGCGACAGACCAAGACTGTCTGTCAAATTAATTAATTAACA | 1045 |
| Db | 91653 | CCACTGTAATCTCAGCTGGGGTGAAGACAGACCAAGACTGTCTCATAAATAATTAATTA | 91712 |
| OY | 1046 | ACGAA 1050 | |
| Db | 91713 | AATAA 91717 | |
| RESULT 13 | | | |
| HS95C20/c | | | |
| LOCUS | HS95C20 | 138849 bp | DNA linear PRI 05-MAR-2003 |
| DEFINITION | Human DNA sequence from clone Rpl-95C20 on chromosome Xp11.3-11.4, complete sequence. | | |
| ACCESSION | Z97181 | | |
| VERSION | Z97181.1 | GI:3059064 | |
| KEYWORDS | HTG. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 138849) | | |
| TITLE | McDougal, R. | | |
| JOURNAL | Direct Submissions | | |
| COMMENT | Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Apr 18, 1998 this sequence version replaced gi:2222679. ----- Genome Center | | |
| | Center: Wellcome Trust Sanger Institute | | |
| | Center code: SC | | |
| | Web site: http://www.sanger.ac.uk | | |
| | Contact: humquerry@sanger.ac.uk | | |

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chix> RPI-95C20 is from the RPI-1 constructed by the group of Pletzer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

```

FEATURES
source
location/Qualifiers
1..138849
/organism="Homo sapiens"
/mo_type="genomic DNA"
/db_xref="RZPD:RPCT704C2095"
/db_xref="taxon:9606"
/chromosome="X"
/map="p11.3-11.4"
/clone="RP1-95C20"
/clone_1b="RPCT-1"

```

| | | | | | | | |
|-----------------------|-------|--------------|----------|------------|-----|--------|--------|
| Query Match | 9.1% | Score | 455 | DB | 9 | Length | 138849 |
| Best Local Similarity | 77.6% | Pred. No. | 4.1e-108 | | | | |
| Matches | 583 | Conservative | 0 | Mismatches | 150 | Indels | 18 |
| | | | | | | Gaps | 2 |

| | | | |
|----|--------|--|--------|
| OY | 333 | TCTATCATCATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAAGCGGAGTCGTCTC | 382 |
| Dd | 122178 | TCGAACATATTATTATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGCGGAGTCGTCTC | 122119 |
| OY | 393 | TGTGCCCACAGCGTGAGATGCAGTGGCGGGATCTGGGCTCACTGCMAAGCTCCGCTCCCCG | 452 |
| Dd | 122118 | TGTGGCCCAAGCGGGGAGTGCAGTGGCGGCAATCTGGCTCACTGMAACTCCGCTCCC GG | 122059 |
| OY | 453 | GTTCAAGCCATCTCTGTGCTGAGCTCCAAAGTAGCTGGGAGCTAAGGGGCCCGGCACCT | 512 |
| Dd | 122058 | GTTCAAGCCATCTCTGTGCTGAGCTCCAAAGTAGCTGGGAGCTAAGGGGCCCGGCACATC | 121999 |
| OY | 513 | ACGCCGGGCTAATTTTTTTTGTATTTTGTAGAGACGGGGTTTCAACGTTTTATGCCGGGA | 572 |
| Dd | 121998 | ACGCCGGGCTAATTTTTTTTGTATTTTGTAGAGACGGGGTTTCAACGTTGTATGCCAGGA | 121939 |
| OY | 573 | TGGCCTCGATCTCTGACCTCGTGATCCGCCGCTCGGCTCCCAAAGTGTGGGATTA | 632 |
| Dd | 121938 | TGGTCTCGATCTCTGACCTCGTGATCCGCCGCTCGGCTCCCAAAGTGTGGGATTA | 121879 |
| OY | 633 | CAGCGTGAGCCACCGCGCGCGGGCCATGATCATCTTCTTGACTATGCTGATGTGACAAGT | 692 |
| Dd | 121878 | CAAGCGTGAGCCACCGCGCGCGGGCCATATTATTCTTTAATAGATTAAGTITCTACA | 121819 |
| OY | 693 | ACCTAAGAOCATCAGACTCTACCTCTTAAATA-----TGCAGTTTGGGCGACGACCGT | 746 |
| Dd | 121818 | TCTAACCTCGATATTTTGGCCGTTCAGTTCTAAACCTACACAGTTTCTCCACGCCGGCGTGT | 121759 |
| OY | 747 | GGCTCATGCTGTAAATTCAGACACTTTGGAGGCGAGAGTGGGTGAATCACTTGAGGCCA | 806 |
| Dd | 121758 | GGCTCATGCTGTAAATTCAGACACTTTGGAGGCGGAGAGTGTGACGATTCACCTGAGGTCA | 121699 |
| OY | 807 | GGAGTTTGAGCACCGCTGGGCCACATGSGTGAACCTGTCTTTTACTTAATAAAAAAAA | 866 |

Db 121698 GGAGTTGGAGACAGCCTGGCCAAATGATGTAACCTGTCTCTACT-----A 121651

Qy 867 AAAAAAAAAATCAGCCGGGTGCTGGGGCACACCTGTATCCAGCTATGCTGGAGG 926

Db 121650 AAAATTCAAAATATAGCTGGGCTGTGGCGGGACCTGTAGTCTCCAGCTCTCAGGAGG 121591

Qy 927 CTGAGGACGAGAGTCACTTGACCCCTGGAGCGGAGGTGACAGTGGCCGAGATCAAT 986

Db 121590 CTGAGGACGAGAGTCACTTGACCCCTGGAGCGGAGGTGACAGTGGCCGAGATGAGCCGAGATGGCC 121531

Qy 987 CACCGCCCTCAGCCTGGGCGACAGACAGCACTCTGTCTCAATTAATTAATAACAAA 1046

Db 121530 CACTGCACTCCAGCCTGGGTGACAGATGAGCACTCTCAATTAATTAATAATAGT 121471

Qy 1047 CGAACAGCAGCTTTGTTTACTTATAGTTATA 1077

Db 121470 AAATTAACATGACAGCTCTGCTTGAATA 121440

RESULT 14

AC114876/c 171697 bp DNA linear PRI 01-JUN-2002

LOCUS Homo sapiens chromosome 3 clone RP11-728B21, complete sequence.

DEFINITION AC114876 AC026918

ACCESSION AC114876.2 GI:21306686

VERSION HTG.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 171697)

AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C., and Haugen, E.D.

TITLE Unpublished

JOURNAL 2 (bases 1 to 171697)

REFERENCE Direct Submission

AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.

JOURNAL Submitted (12-MAR-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

REFERENCE 3 (bases 1 to 171697)

AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C., and Haugen, E.D.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

REFERENCE On Jun 1, 2002 this sequence version replaced gi:19352302.

COMMENT ----- Genome Center

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: uwgchgs@u.washington.edu

Drafting Center: BCM

----- Project Information

Center Project name: chr-3

Center Clone name: RP11-728B21 (bc0540)

----- Summary Statistics

Sequencing vector: unknown; 53% of reads

Sequencing vector: plasmid; 47% of reads

Chemistry: Dye-terminator ET; 89% of reads

Chemistry: Dye-terminator Big Dye; 11% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 171547 bases at least Q40

Consensus quality: 171691 bases at least Q30

Consensus quality: 171697 bases at least Q20

Insert size: 171697; sum-of-contigs

Quality coverage: 9.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-349B16 (UWGC:bc0646) AC098481

3': RP11-624D20 (UWGC:bc0494) AC068222

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

| SeqDerMap | Fingerprint | SeqDerMap | Fingerprint | SeqDerMap | Fingerprint |
|-----------|-------------|-----------|-------------|-----------|-------------|
| 2292 | 2308 | 6857 | 6790 | 8636 | 8687 |
| 2067 | 2005 | 6382 | 6450 | 6 | <800 |
| 4763 | 4915 | 512 | <800 | 3403 | 3464 |
| 2585 | 2578 | 449 | <800 | 334 | <800 |
| 1981 | 2005 | 7940 | 7924 | 951 | 930 |
| 5496 | 5398 | 10618 | 10372 | 4919 | 4873 |
| 8296 | 8303 | 305 | <800 | 3694 | 3709 |
| 10488 | 10186 | 1052 | 1054 | 5455 | 5487 |
| 5655 | 5680 | 2795 | 2867 | 555 | <800 |
| 6535 | 6574 | 1527 | 1521 | 13221 | 13066 |
| 2695 | 2731 | 9418 | 9389 | 8640 | 8687 |
| 10234 | 10186 | 2970 | 3015 | 5654 | 5681 |
| 4871 | 4915 | 10593 | 10372 | 2842 | 2844 |
| 1822 | 1781 | 891 | 883 | 10800 | 10593 |
| 726 | 720 | 1004 | 1054 | 11874 | 11745 |
| 540 | <800 | 1340 | 1314 | 8529 | 8687 |
| 11711 | 11427 | 817 | 883 | 286 | <800 |
| 25420 | 25600 | 4767 | 4687 | 6660 | 6724 |

| | | | | | |
|------|------|-------|-------|-------|-------|
| 2222 | 2308 | 7571 | 7564 | 13999 | 13754 |
| 360 | <800 | 4492 | 4508 | 9092 | 9112 |
| 2344 | 2308 | 1092 | 1054 | 2356 | 2380 |
| 1198 | 1180 | 3332 | 3396 | 2287 | 2380 |
| 3828 | 3831 | 1123 | 1054 | 2669 | 2713 |
| 3328 | 3300 | 1803 | 1824 | 924 | 930 |
| 3248 | 3300 | 3041 | 3122 | 4641 | 4650 |
| 1289 | 1252 | 16514 | 16912 | 1247 | 1206 |
| 399 | <800 | 192 | <800 | 1708 | 1662 |
| 4134 | 4173 | 4763 | 4687 | 6082 | 6133 |
| 900 | 895 | 587 | <800 | 1904 | 1897 |
| 3852 | 3963 | 5086 | 5041 | 4433 | 4394 |
| 4463 | 4378 | 2278 | 2421 | 6851 | 6941 |
| 3959 | 3831 | 1336 | 1314 | 1799 | 1783 |
| 1990 | 2005 | 3223 | 3122 | 23865 | 24086 |
| 4272 | 4173 | 7004 | 7057 | | |
| 2327 | 2308 | 971 | 1054 | | |
| 6858 | 6866 | 2341 | 2421 | | |
| 3056 | 3063 | 1649 | 1646 | | |
| 6815 | 6866 | 2822 | 2867 | | |
| 809 | 810 | 6084 | 6048 | | |
| 4433 | 4378 | 163 | <800 | | |
| 6135 | 6032 | 9648 | 9389 | | |
| | | 311 | <800 | | |
| | | 3279 | 3396 | | |
| | | 3900 | 3950 | | |
| | | 2035 | 2029 | | |
| | | 3612 | 3630 | | |
| | | 3053 | 3229 | | |
| | | 3414 | 3396 | | |
| | | 865 | 883 | | |
| | | 637 | <800 | | |
| | | 259 | <800 | | |
| | | 97 | <800 | | |
| | | 1242 | 1314 | | |
| | | 270 | <800 | | |
| | | 23 | <800 | | |

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 227137)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (23-OCT-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 227137)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (15-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 15, 2002 this sequence version replaced gl:16328287.

Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Project Information
Center clone name: Rpl1-349E16 (bc0646)

Summary Statistics
Sequencing Vector: plasmid, L08752, 100% of reads
Chemistry: Dye-terminator ET; 76% of reads
Chemistry: Dye-terminator Big Dye; 24% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 227062 bases at least Q40
Consensus quality: 227137 bases at least Q30
Consensus quality: 227137 bases at least Q20
Insert size: 227137; sum-of-contigs
Quality coverage: 8.0x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': Rpl1-728B21 (UWGC:bc0540) AC114876, 123790-bp overlap
3': Rpl1-286L5 (UWGC:bc0344) AC103559

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

NSII
EcoRI
HindIII
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

3221 3262 3904 3909 305 <800

| | | | | | |
|-------|-------|-------|-------|-------|-------|
| 22527 | 22843 | 8696 | 8633 | 3211 | 3196 |
| 7364 | 7368 | 6 | <800 | 6382 | 6483 |
| 1250 | 1186 | 4641 | 4671 | 512 | <800 |
| 1646 | 1681 | 924 | 937 | 449 | <800 |
| 5862 | 5809 | 2669 | 2702 | 1388 | 1337 |
| 4415 | 4376 | 2286 | 2141 | 1649 | 1684 |
| 5203 | 5251 | 2356 | 2346 | 2341 | 2357 |
| 2339 | 2380 | 9092 | 9731 | 971 | 998 |
| 4104 | 4145 | 13999 | 13681 | 7003 | 7067 |
| 2067 | 2088 | 6660 | 6723 | 3223 | 3196 |
| 1413 | 1416 | 286 | <800 | 1336 | 1337 |
| 8832 | 8730 | 8529 | 8633 | 2278 | 2251 |
| 1451 | 1416 | 11874 | 11716 | 5086 | 5023 |
| 3108 | 3003 | 10800 | 10477 | 587 | <800 |
| 1155 | 1088 | 2842 | 2850 | 4763 | 4697 |
| 10132 | 9934 | 5654 | 5704 | 192 | <800 |
| 7178 | 7368 | 8640 | 8633 | 16514 | 16792 |
| 2392 | 2380 | 13221 | 12924 | 3041 | 3013 |
| 4819 | 4786 | 555 | <800 | 1803 | 1813 |
| 1188 | 1186 | 5455 | 5489 | 1123 | 1071 |
| 2666 | 2689 | 3694 | 3751 | 3332 | 3382 |
| 1188 | 1186 | 4919 | 4887 | 1092 | 1071 |
| 431 | <800 | 951 | 937 | 4492 | 4528 |
| 732 | 750 | 334 | <800 | 7571 | 7668 |
| 1680 | 1681 | 3403 | 3506 | 4767 | 4697 |
| 1855 | 1877 | 3587 | 3633 | 817 | 832 |
| 1537 | 1490 | 2819 | 2850 | 1340 | 1327 |
| 1361 | 1350 | 3838 | 3909 | 1004 | 998 |
| 5212 | 5251 | 5684 | 5704 | 891 | 922 |
| 24 | <800 | 373 | <800 | 10593 | 10357 |
| 1608 | 1681 | 63 | <800 | 2970 | 2846 |
| 1731 | 1877 | 203 | <800 | 9418 | 9439 |
| 1439 | 1416 | 7896 | 7870 | 1527 | 1526 |
| 327 | <800 | 332 | <800 | 2795 | 2775 |
| 5312 | 5251 | 407 | <800 | 1052 | 1071 |
| 3221 | 3262 | 3904 | 3909 | 305 | <800 |

| | | | | | |
|-------|-------|-------|-------|-------|-------|
| 7582 | 7368 | 3318 | 3369 | 10618 | 10357 |
| 61 | <800 | 44 | <800 | 7321 | 7358 |
| 1886 | 1877 | 2371 | 2346 | 2226 | 2251 |
| 1453 | 1416 | 6367 | 6431 | 1370 | 1327 |
| 4192 | 4145 | 3577 | 3633 | 5827 | 5912 |
| 17721 | 17419 | 303 | <800 | 371 | <8000 |
| 7973 | 7948 | 2574 | 2596 | 9765 | 9439 |
| 395 | <800 | 629 | <800 | 5615 | 5530 |
| 2993 | 3003 | 2820 | 2850 | 7814 | 7838 |
| 646 | <800 | 9059 | 9045 | 1345 | 1327 |
| 3902 | 3887 | 1553 | 1517 | 978 | 998 |
| 74 | <800 | 1792 | 1741 | 2723 | 2775 |
| 1930 | 2088 | 7744 | 7870 | 1059 | 1071 |
| 73 | <800 | 18454 | 18910 | 2640 | 2595 |
| 5859 | 5809 | 4671 | 4671 | 4019 | 4073 |
| 1198 | 1186 | 8958 | 9045 | 3118 | 3126 |
| 4305 | 4376 | | | 5282 | 5530 |
| 6273 | 6405 | | | 83 | <8000 |
| 1719 | 1681 | | | 3648 | 3652 |
| 482 | <800 | | | 5541 | 5530 |
| 504 | <800 | | | 432 | <8000 |

| | | | | |
|-----------------------|--------------|---------------------|-----------------|----------------|
| Query Match | 9.1%; | Score 455; | DB 9; | Length 227137; |
| Best Local Similarity | 78.9%; | Pred. No. 4.4e-108; | | |
| Matches 572; | Conservative | 0; | Mismatches 140; | Indels 13; |
| | | | | Gaps 22; |

| | | | |
|----|-------|---|-------|
| QY | 326 | GGTGGTGCTAATGATGATGATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGT | 385 |
| Db | 91443 | GATGGTTCATTTCTCCAGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGATGACCGAGT | 91502 |
| QY | 386 | CTCGCTCTGTGCGCCAGGCTGGAGTGGAGTGGCGGGATCTCGGCTCATCTGACGCTCCG | 445 |
| Db | 91503 | CTTGCTCTGTGGCCCAAGTGGAGTGCATGGGCGCAATCTCGGCTCATCTGACGCTCCG | 91562 |
| QY | 446 | CTCCGGGTTTCAGCGCATTTCTCTGCTCAGCCTCCCAAGTACGTGGACCTACAGCGCG | 505 |
| Db | 91563 | CTCCAGGTTTCAGCGCATTTCTCTGCTCAGCCTCCCAAGTACGTGGACCTACAGCGCG | 91622 |
| QY | 506 | CGCCACTACGCGCGGCTAATTTTTTTGTATTTTTTAGTAGAGCGGGTTTCAACGTTTTA | 565 |
| Db | 91623 | CGCCACACGCGCGGCTAATTTTTTTGTATTTTTTAGTAGAGCGGGTTTCAACGTTTTA | 91682 |
| QY | 566 | GC CGGGAATGGCTGATCTCTGACCTCGATGATCCGCGCGGCTCGGCTCCCAAGTGC | 625 |
| Db | 91683 | GCCAGATGGTCTGATCTCTGATCTCTGACTCTCGATGATCCGCGCGGCTCGGCTCCCAAGTGC | 91742 |
| QY | 626 | GGGATTACAGGCGTGAAGCACAACGCGCCCGGCGCATGATCTTCTGACTATGCTGATGT | 685 |
| Db | 91743 | GGGATTACAGGCGTGAAGCACAACGCGCCCGGCGCATGATCTTCTGACTATGCTGATGT | 91802 |
| QY | 686 | GACAAATTAAGTAAAGCATCACTTAACTTTAAATATGACGTTTGGCGCAGGACCG | 745 |

| | | | |
|----|-------|--|-------|
| Db | 91802 | GCTCATCTAATGAAGAAATTATGACAGGGGTGTTAAAACTAGGACGGGGATGGGATGA | 91861 |
| QY | 746 | TGGCTCATGCTGTAAATTCAGACACTTTGGGAGGCAAGAGTGGGTGAATCATTGAGGCC | 805 |
| Db | 91862 | TGGCTCAGCGCTGTAAATCCAGACACTTTGGGAGGCCAAGGCCAGTGGATCACCCTGAGTTC | 91921 |
| QY | 806 | AGGAGTTTGAGACCAGCGCTGGCCCAATGGTGAACCTGTCTTTACTTAAAAA | 865 |
| Db | 91922 | CGGAGTTTGGAGACCAGTCTGGCCAAAGTGCMAAACCCACTCTACT----- | 91969 |
| QY | 866 | AAAAAAAAAAAAATCAGCCGGGTGTGCTGGGGCACACTGTATATCCCACTATGCTGGAG | 925 |
| Db | 91970 | AAAAATCAAAAAATTAATCCAGGGCTTGTGTGTGTGTCTGTATCCCACTACTCGGGAG | 92029 |
| QY | 926 | GCTGAGGACGAGAGTCACTTGAACCTGGAGCGGAGGTGCAGTGGGCCAGATCACA | 985 |
| Db | 92030 | TCTGAGGCAATGAGAAATACCTTGAACCTGGGAGGTGAGAGTTGCATGAGCCAAAGATTACG | 92089 |
| QY | 986 | TCACCGGCTTCAGACCTGGGCGACAGAGCAAGACTGTCTCAATTAATAATAACA | 1045 |
| Db | 92090 | CCACTGTACTCCAGCGCTGGGTGACAGAGCAAGACTCTGTCTCAAAAAATAATAATAA | 92149 |
| QY | 1046 | ACGAA 1050 | |
| Db | 92150 | AATAA 92154 | |

Search completed: February 26, 2005, 06:23:44
Job time : 21330.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 / Search time 2531.5 Seconds

(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271B-13_COPY_1_5000

Perfect score: 5000

Sequence: 1 gggccagagtgaggaag.....tgcatccacagtgctgctc 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1980s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 5000 | 100.0 | 11570 | 3 | AAA95905 Human KLK |
| 2 | 5000 | 100.0 | 11570 | 12 | ADK52482 Human KAL |
| 3 | 5000 | 100.0 | 11570 | 13 | ADR72623 Human KAL |
| 4 | 5000 | 100.0 | 11570 | 13 | ADR72623 Human KAL |
| 5 | 472 | 9.4 | 525 | 12 | ACH80345 Human KAL |
| 6 | 440.2 | 8.8 | 41454 | 11 | ACN44942 Human KAL |
| 7 | 424.4 | 8.5 | 84607 | 2 | AAK90847 Human PAC |
| 8 | 419 | 8.4 | 91352 | 12 | ADN94799 Human PAC |
| 9 | 418.4 | 8.4 | 115935 | 13 | ADR52890 Drug Ther |
| 10 | 417.6 | 8.4 | 26371 | 11 | ACN44814 Human gen |
| 11 | 417.6 | 8.4 | 91760 | 11 | ACN44814 Human gen |
| 12 | 417.4 | 8.3 | 216215 | 10 | ADK59167 Human MP5 |
| 13 | 416.6 | 8.3 | 139904 | 6 | ABK83562 Human CDN |
| 14 | 416.2 | 8.3 | 79652 | 10 | ADG70184_3 |
| 15 | 415.2 | 8.3 | 17397 | 4 | AAK36445 Human car |
| 16 | 415.2 | 8.3 | 17397 | 10 | ADK47139 Human car |
| 17 | 415.2 | 8.3 | 17397 | 13 | ADJ08557 Human car |
| 18 | 415.2 | 8.3 | 19334 | 4 | AAK36443 Human car |
| 19 | 415.2 | 8.3 | 19334 | 10 | ADK47137 Human car |
| 20 | 415.2 | 8.3 | 19334 | 13 | ADJ08555 Human car |

| | | | | | |
|----|-------|-----|--------|----|----------|
| 21 | 415.2 | 8.3 | 19345 | 4 | AAK36444 |
| 22 | 415.2 | 8.3 | 19345 | 10 | ADK47138 |
| 23 | 415.2 | 8.3 | 19345 | 13 | ADJ08556 |
| 24 | 415.2 | 8.3 | 29001 | 12 | ADK85575 |
| 25 | 414.8 | 8.3 | 12970 | 4 | AAK05001 |
| 26 | 414.8 | 8.3 | 12970 | 4 | ABL7894 |
| 27 | 414.6 | 8.3 | 74037 | 6 | ABK94412 |
| 28 | 414.6 | 8.3 | 11084 | 12 | ADJ08608 |
| 29 | 414.2 | 8.3 | 212231 | 11 | ACN44598 |
| 30 | 410.4 | 8.2 | 10445 | 4 | AAK03652 |
| 31 | 409.6 | 8.2 | 50000 | 9 | ADK16926 |
| 32 | 408.2 | 8.2 | 240000 | 8 | ACD13446 |
| 33 | 407.6 | 8.2 | 196686 | 11 | ACN44170 |
| 34 | 407 | 8.1 | 32460 | 12 | ADK07738 |
| 35 | 406.2 | 8.1 | 12970 | 4 | AAK05001 |
| 36 | 406.2 | 8.1 | 12970 | 4 | ABL7894 |
| 37 | 406 | 8.1 | 32224 | 4 | AAK89966 |
| 38 | 405.8 | 8.1 | 4433 | 4 | AAK04142 |
| 39 | 405.8 | 8.1 | 4433 | 5 | ABK18400 |
| 40 | 405.8 | 8.1 | 4433 | 5 | AAK40516 |
| 41 | 405.8 | 8.1 | 4433 | 11 | ADK09722 |
| 42 | 405.8 | 8.1 | 126001 | 12 | ADK77123 |
| 43 | 405 | 8.1 | 32767 | 8 | AAK56075 |
| 44 | 405 | 8.1 | 32767 | 9 | ADK02437 |
| 45 | 405 | 8.1 | 32767 | 10 | ADK72176 |

ALIGNMENTS

| | | |
|----------|--|--------------------------|
| RESULT 1 | AAA95905 | standard; DNA; 11570 BP. |
| ID | AAA95905 | |
| XX | AAA95905; | |
| AC | 02-FEB-2001 | (first entry) |
| DT | XX | |
| XX | XX | |
| DE | Human KLK-L2 gene. | |
| XX | XX | |
| KW | Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; | |
| KM | Kallikrein-like protein; serine protease; cytosolic; cancer; | |
| KM | prostate cancer; ds. | |
| XX | XX | |
| OS | Homo sapiens. | |
| XX | XX | |
| PN | WO200053776-A2. | |
| XX | XX | |
| PD | 14-SEP-2000. | |
| XX | XX | |
| PF | 09-MAR-2000; 2000WO-CO00258. | |
| XX | XX | |
| PR | 11-MAR-1999; 99US-0124260P. | |
| PR | 01-APR-1999; 99US-0127386P. | |
| PR | 21-JUL-1999; 99US-0144919P. | |
| PA | (MOUN) MOUNT SINAI HOSPITAL. | |
| PI | Yousef GM, Diamandis EP; | |
| PI | WPI, 2000-587440/55. | |
| DR | P-FSDB; AAK51296. | |
| XX | XX | |
| PT | New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L | |
| XX | protein mediated disorders, especially cancer. | |
| PS | Claim 1; Page 143-149; 184pp; English. | |
| XX | XX | |
| CC | The present sequence is the coding sequence of the human KLK-L2 gene, | |
| CC | which encodes a kallikrein-like protein. Kallikreins and kallikrein-like | |
| CC | proteins are a subgroup of the serine protease enzyme family. They | |
| CC | catalyse the selective cleavage of specific polypeptide precursors to | |
| CC | release peptides with potent biological activity. Nucleic acids encoding | |

| | |
|----------|-----------|
| AAK36444 | Human car |
| ADK47138 | Human car |
| ADJ08556 | Human car |
| ADK85575 | Human zln |
| AAK05001 | Human rep |
| ABL7894 | Human tes |
| ABK94412 | Human tes |
| ADJ08608 | Human sof |
| ACN44598 | Human gen |
| AAK03652 | Human rep |
| ADK16926 | Human DYX |
| ACD13446 | Human DNA |
| ACN44170 | Human gen |
| ADK07738 | Nucleotid |
| AAK05001 | Human rep |
| ABL7894 | Human tes |
| AAK89966 | Human dig |
| AAK04142 | Human rep |
| ABK18400 | Human rep |
| AAK40516 | Human rep |
| ADK09722 | Human pro |
| ADK77123 | Human PAZ |
| AAK56075 | Human SNL |
| ADK02437 | Human SNL |
| ADK72176 | Human SNL |

CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 5000; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGGCCAGAGTGAAGGCAAGAGAGGTTGAGAGCTCCCTGTCAAAAGTGCTTGAAGTC 60
DB 1 GGGCCAGAGTGAAGGCAAGAGAGGTTGAGAGCTCCCTGTCAAAAGTGCTTGAAGTC 60
QY 61 TCCCTGCTTAAATGCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 TCCCTGCTTAAATGCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 ACACAAATGAGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACAC 300
DB 241 ACACAAATGAGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACAC 300
QY 301 AAAGGGGACGAAGCAGTTTTCAGGGTGCTGATCAATCACTTTTCTTTTCTTTTCTTT 360
DB 301 AAAGGGGACGAAGCAGTTTTCAGGGTGCTGATCAATCACTTTTCTTTTCTTTTCTTT 360
QY 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 420
DB 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 420
QY 421 GATCTGGGCTCACTGCAAGCTCCGCTCCGGGTTACACGCATTTCTTGGCTCAAGCTTC 480
DB 421 GATCTGGGCTCACTGCAAGCTCCGCTCCGGGTTACACGCATTTCTTGGCTCAAGCTTC 480
QY 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCGCTAATTTTGTATTTTAA 540
DB 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCGCTAATTTTGTATTTTAA 540
QY 541 GTAGAGACGGGTTTCAACGTTTAAACCGGATGAGCTGATCTCTGACCTGTGATCC 600
DB 541 GTAGAGACGGGTTTCAACGTTTAAACCGGATGAGCTGATCTCTGACCTGTGATCC 600
QY 601 GCCCGCTGGGCTCCCAAAAGTGTGGATTACAGGCGGTGAGCCACCGGCGCCAGT 660
DB 601 GCCCGCTGGGCTCCCAAAAGTGTGGATTACAGGCGGTGAGCCACCGGCGCCAGT 660
QY 661 ATCATCTTCTGACTAGTGTGATGATGACAAATGACCTAAAGCCATGACCTTACCTTTA 720
DB 661 ATCATCTTCTGACTAGTGTGATGATGACAAATGACCTAAAGCCATGACCTTACCTTTA 720
QY 721 AATATGACGTTTGGGCAAGGACCGTGCTCATGCTGTAAATTCAGACCTTTGGAGGC 780
DB 721 AATATGACGTTTGGGCAAGGACCGTGCTCATGCTGTAAATTCAGACCTTTGGAGGC 780
QY 781 AAGAGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCAGCTGGCCAAATGTGAAA 840
DB 781 AAGAGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCAGCTGGCCAAATGTGAAA 840
QY 841 CTCTGTCTTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
DB 841 CTCTGTCTTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 900
```

```
QY 901 ACTGTAAATCCAGCATATGCTGAGGCTGAGGACAGAGATCATTTGAACCTTGAGGCG 960
DB 901 ACTGTAAATCCAGCATATGCTGAGGCTGAGGACAGAGATCATTTGAACCTTGAGGCG 960
QY 961 GAGGTTGACGTGGCCGAGATCATACCGCTTCACGCTGGGCGACAGACAGACT 1020
DB 961 GAGGTTGACGTGGCCGAGATCATACCGCTTCACGCTGGGCGACAGACAGACT 1020
QY 1021 CTGTCTCAATTAATAATAAATAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 CTGTCTCAATTAATAATAAATAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 AAAAAAAAAATGCTGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 AAAAAAAAAATGCTGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 AGAATCTTAAGTATTTTGAAGAAATCATTCAGAACTTTTAAAAAGAAATCACAGA 1200
DB 1141 AGAATCTTAAGTATTTTGAAGAAATCATTCAGAACTTTTAAAAAGAAATCACAGA 1200
QY 1201 GGCATTGAAAGACAGGAGAGAGACAGGAGAGACAGAAACACTGTGGCCCAAGGAGACA 1260
DB 1201 GGCATTGAAAGACAGGAGAGAGACAGGAGAGACAGAAACACTGTGGCCCAAGGAGACA 1260
QY 1261 ACAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 ACAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCCAGAGGGATTAAAGA 1380
DB 1321 AAAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCCAGAGGGATTAAAGA 1380
QY 1381 GAGAGGGGGTGGAGAGAGACAGAGATTGAGAGAGAGCTGAGAAAGTAGCCGAGGGA 1440
DB 1381 GAGAGGGGGTGGAGAGAGACAGAGATTGAGAGAGAGCTGAGAAAGTAGCCGAGGGA 1440
QY 1441 GAACCAAGAGAGATGAGAAAGAGCTGTGAGAAAAACAGAGACAAAGATGAAAGAGG 1500
DB 1441 GAACCAAGAGAGATGAGAAAGAGCTGTGAGAAAAACAGAGACAAAGATGAAAGAGG 1500
QY 1501 AGTATGAGGGTGAACAGACAGTGTGGAATGAGCAAAATGCAAGAAAGCAAGCAA 1560
DB 1501 AGTATGAGGGTGAACAGACAGTGTGGAATGAGCAAAATGCAAGAAAGCAAGCAA 1560
QY 1561 TCCAGGCGCAAGAAATGAGAACCCAGAGTTGTTGAGAAAGCCAGATCTTAAAGCTGGGG 1620
DB 1561 TCCAGGCGCAAGAAATGAGAACCCAGAGTTGTTGAGAAAGCCAGATCTTAAAGCTGGGG 1620
QY 1621 AAGCAGGGAAGGGGCTGGCTTCCGAGAGCCCTCCCATTTCTCCGGGCAAGGAG 1680
DB 1621 AAGCAGGGAAGGGGCTGGCTTCCGAGAGCCCTCCCATTTCTCCGGGCAAGGAG 1680
QY 1681 GTAGGAGTGAATTCGGAATGGGTTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
DB 1681 GTAGGAGTGAATTCGGAATGGGTTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
QY 1741 GTAGGAGTGAATTCGGAATGGGTTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1800
DB 1741 GTAGGAGTGAATTCGGAATGGGTTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1800
QY 1801 CAGTGGGTGTTAATACTAGGCTGGGCTGGGCTCCAGAGCCAGAGAGAGGAGTGGCCAGGA 1860
DB 1801 CAGTGGGTGTTAATACTAGGCTGGGCTGGGCTCCAGAGCCAGAGAGAGGAGTGGCCAGGA 1860
QY 1861 AAGCAAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCCACCTCTACCTGG 1920
DB 1861 AAGCAAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCCACCTCTACCTGG 1920
QY 1921 GGGAGAGGCAAGTGAAGCTGTGTGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGTGTCT 1980
DB 1921 GGGAGAGGCAAGTGAAGCTGTGTGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGTGTCT 1980
QY 1981 GTGCGTCTGACCAACATCTTTCTGTGCTCCCTCTGTGCTGTGAGAGCTGTAGA 2040
```

1981 GTGGGCTCTGCAACCAATCTTTCTGTGTCCCTCTTGCCCTGTGTGAGGCTGTAGA 2040
QY CTCTATCTTCTGAAATCTATATGTCCTGGGTCTCAGCGCAGATGCCATGTGTGCCCTGC 2100
Db CTCTATCTTCTGAAATCTATATGTCCTGGGTCTCAGCGCAGATGCCATGTGTGCCCTGC 2100
QY CTGTGGTTCCTCTCTACCTGGGGAATTAAGTATGGGAGGAGGAGGAGGAGGAGGAGG 2160
Db CTGTGGTTCCTCTCTACCTGGGGAATTAAGTATGGGAGGAGGAGGAGGAGGAGGAGG 2160
QY GCTCCCGGATCGCTGGGCTCCCAACCCCTGTGACATTCCTCATCCAGGTGACGCGCC 2220
Db GCTCCCGGATCGCTGGGCTCCCAACCCCTGTGACATTCCTCATCCAGGTGACGCGCC 2220
QY ATGGCTACAGCAAAACCCCTGTGATGGGTCTGTGTCTGTGATCAAGCTTGTCTT 2280
Db ATGGCTACAGCAAAACCCCTGTGATGGGTCTGTGTCTGTGATCAAGCTTGTCTT 2280
QY CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
QY CTCTGCGGCACTAGAGCGCTGTCCCTGGGGAATGTGTGAGCGCTGGGCACTGACCTCG 2400
Db CTCTGCGGCACTAGAGCGCTGTCCCTGGGGAATGTGTGAGCGCTGGGCACTGACCTCG 2400
QY GATGGGGTGAATGATGATCTGTCTGTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db GATGGGGTGAATGATGATCTGTCTGTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY GACTGCCACGGGT 2520
Db GACTGCCACGGGT 2520
QY CAGGTGGCACTGACCTTTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db CAGGTGGCACTGACCTTTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY ATTGT 2640
Db ATTGT 2640
QY GT 2700
Db GT 2700
QY TCCCTGAGGTCCTCCGGGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db TCCCTGAGGTCCTCCGGGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
QY GT 2820
Db GT 2820
QY TTTGT 2880
Db TTTGT 2880
QY TCTGT 2940
Db TCTGT 2940
QY TTTGT 3000
Db TTTGT 3000
QY CAGGTGACTGACTGT 3060
Db CAGGTGACTGACTGT 3060
QY TGTGCAACAACGCGATCTGTGCGTGTGCACTGAGACACTGTGTGTGTGTGTGTGTGTGT 3120
Db TGTGCAACAACGCGATCTGTGCGTGTGCACTGAGACACTGTGTGTGTGTGTGTGTGTGT 3120

3061 TGTGCAACAACGCGATCTGTGCGTGTGCACTGAGACACTGTGTGTGTGTGTGTGTGTGT 3120
QY GTTAGGCTGCCCGGAGCGGTGTGTACTGTGAGACAGAGCTGTATGTAGCTGTGACCTGTG 3180
Db GTTAGGCTGCCCGGAGCGGTGTGTACTGTGAGACAGAGCTGTATGTAGCTGTGACCTGTG 3180
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
QY GCGTGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Db GCGTGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
QY GCGTGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
Db GCGTGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3660
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3660
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3720
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3720
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3840
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3840
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3900
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4020
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4020
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
QY GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200
Db GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200

| | | | |
|----|------|---|------|
| OY | 4201 | CCAGGTATCCGCCCACTCGGACCTCCCAAGTGCTGGGGTTTACAAGCTAAGCACCCG | 4200 |
| Db | 4201 | TCAGGTGATCCGCCCACTCGGACCTCCCAAGTGCTGGGGTTTACAAGGCTAAGCACCCG | 4200 |
| OY | 4261 | CCCGAGCCCAAGTCAGAGCTCTTATATGAGAGACTTAACTATGAACCTTGACCCTGGAC | 4320 |
| Db | 4261 | CCCGAGCCCAAGTCAGAGCTCTTATATGAGAGACTTAACTATGAACCTTGACCCTGGAC | 4320 |
| OY | 4321 | CTAACTAAGTCAATTCCAAACCCCTTCTGCTCCAGCCCTGAGCCCACTCACTGAGGCC | 4380 |
| Db | 4321 | CTAACTAAGTCAATTCCAAACCCCTTCTGCTCCAGCCCTGAGCCCACTCACTGAGGCC | 4380 |
| OY | 4381 | TGACCCCACTCTTGAGAGCAGATTCATCCCTAAAGCCCTGCTCCCTCCCATCCCCAG | 4440 |
| Db | 4381 | TGACCCCACTCTTGAGAGCAGATTCATCCCTAAAGCCCTGCTCCCTCCCATCCCCAG | 4440 |
| OY | 4441 | GCTCCAGCCCCCAcAGCTTTTGgCACTAcccCTGAGCTTGTCCAGGAATCTGTACCCAAAT | 4500 |
| Db | 4441 | GCTCCAGCCCCCAcAGCTTTTGgCACTAcccCTGAGCTTGTCCAGGAATCTGTACCCAAAT | 4500 |
| OY | 4501 | TTTAAcCTTCaATGTATgTTCTTAAGCCAAATTCAGGAATCTGTAGAGTCCAGTTAAATTCcA | 4560 |
| Db | 4501 | TTTAAcCTTCaATGTATgTTCTTAAGCCAAATTCAGGAATCTGTAGAGTCCAGTTAAATTCcA | 4560 |
| OY | 4561 | GTAACCTCACTGAGCCTGGGCTGTCTTGTAGAGCTTGAGCCTGAGGCTTGAGAGGTGCA | 4620 |
| Db | 4561 | GTAACCTCACTGAGCCTGGGCTGTCTTGTAGAGCTTGAGCCTGAGGCTTGAGAGGTGCA | 4620 |
| OY | 4621 | CTCTTATTTCTCAAGGcCTGTGccCCTGTGccCCTCAcAGATGTCAgACACCCACCTCTTAAGCT | 4680 |
| Db | 4621 | CTCTTATTTCTCAAGGcCTGTGccCCTGTGccCCTCAcAGATGTCAgACACCCACCTCTTAAGCT | 4680 |
| OY | 4681 | GGTCGGGcCTGTGAAGTGTGAAGCCCAcCCCAAGCCCAAGCCCGGcCTTAAGCCCGGC | 4740 |
| Db | 4681 | GGTCGGGcCTGTGAAGTGTGAAGCCCAcCCCAAGCCCAAGCCCGGcCTTAAGCCCGGC | 4740 |
| OY | 4741 | CAACCCATTTTTCCGTTCCCAgAGCATGTCTTCGCAACAAATGATGTTTCTGTGACCAcC | 4800 |
| Db | 4741 | CAACCCATTTTTCCGTTCCCAgAGCATGTCTTCGCAACAAATGATGTTTCTGTGACCAcC | 4800 |
| OY | 4801 | CCTCTTAACAacCGTGTGcCTTCTTGgAGCAACCAAGATCTTGgAGCTGGgGcCCGGgAAAGAG | 4860 |
| Db | 4801 | CCTCTTAACAacCGTGTGcCTTCTTGgAGCAACCAAGATCTTGgAGCTGGgGcCCGGgAAAGAG | 4860 |
| OY | 4861 | CCCGGTCCGATGAACAGCAGACAGCCGCAATCAATGATGATTCGgACTGCGCATATGCAACCC | 4920 |
| Db | 4861 | CCCGGTCCGATGAACAGCAGACAGCCGCAATCAATGATGATTCGgACTGCGCATATGCAACCC | 4920 |
| OY | 4921 | AGCCGTGACAGGcCGGcGTGTGTCTTAAGGCCCAACCAcAGCTTACTTGcCGGcCGGTGTGG | 4980 |
| Db | 4921 | AGCCGTGACAGGcCGGcGTGTGTCTTAAGGCCCAACCAcAGCTTACTTGcCGGcCGGTGTGG | 4980 |
| OY | 4981 | TGCATCCCAcAGTGGTGTCTC 5000 | |
| Db | 4981 | TGCATCCCAcAGTGGTGTCTC 5000 | |

| | |
|----------|---|
| RESULT 2 | |
| ADK52482 | |
| ID | ADK52482 standard; DNA, 11570 BP. |
| XX | |
| AC | ADK52482; |
| XX | |
| DT | 03-JUN-2004 (first entry) |
| XX | |
| DE | Human kallikrein 5 encoding sequence. |
| XX | |
| KW | kallikrein 5; cancer; Cytostatic; Immunostimulant; ds |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO2004021008-A2. |

| XX | PD | 11-MAR-2004. |
|----|----|--|
| XX | PF | 28-AUG-2003; 2003WO-CA001310. |
| XX | PR | 28-AUG-2002; 2002US-0407333P. |
| XX | PA | (MOUN) MOUNT SINAI HOSPITAL. |
| XX | PA | (YOUS) YOUSEF G. |
| XX | PI | Diamonds EP; |
| XX | DR | WPI; 2004-239232/22. |
| XX | DR | P-PEDB; ADK52481. |
| PT | PT | Detecting kallikrein 5 associated with cancer, useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma, comprises comparing the detected amount of kallikrein 5 in a sample with a standard sample. |
| PS | PS | Disclosure; SEQ ID NO 2; 59pp; English. |
| CC | CC | The present inventin relates to detecting kallikrein 5 associated with breast or ovarian cancer in a patient comprises detecting in the sample kallikrein 5 and comparing the detected amount with an amount detected for a standard. The method is useful in detecting kallikrein 5 associated with breast or ovarian cancer in a patient. The methods and kits are useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for preventing and treating breast and ovarian cancer and for stimulating or enhancing antibody production or for inducing an immune response. The present sequence represents human kallikrein 5 encoding sequence. |
| CC | CC | Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other; |
| XX | XX | Query Match 100.0%; Score 5000; DB 12; Length 11570; |
| XX | XX | Best Local Similarity 100.0%; Pred. No. 0; |
| XX | XX | Matches 5000; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0 |
| QY | QY | 1 GGGCCCAAGTGAAGCAAGAAAGAGTTGAGAGCTCCCTCTGCAAAAGTGGCTTGAGTC 60 |
| DB | DB | 1 GGGCCCAAGTGAAGCAAGAAAGAGTTGAGAGCTCCCTCTGCAAAAGTGGCTTGAGTC 60 |
| QY | QY | 61 TCCTCTGCTTAAATGCAAGGAGAGGCGCAAAAGACAGGGAAGGAGAGGGGTGGG 120 |
| DB | DB | 61 TCCTCTGCTTAAATGCAAGGAGAGGCGCAAAAGACAGGGAAGGAGAGGGGTGGG 120 |
| QY | QY | 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACAATCAAGAAACAGAGAGAAAC 180 |
| DB | DB | 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAACAATCAAGAAACAGAGAGAAAC 180 |
| QY | QY | 181 ACAGAGAGCTGGGACACAGAGGACACACAGAGTCAAGAGAAAGAGAAATAGAGAAAG 240 |
| DB | DB | 181 ACAGAGAGCTGGGACACAGAGGACACACAGAGTCAAGAGAAAGAGAAATAGAGAAAG 240 |
| QY | QY | 241 ACACAAATGAGACACAGAGGTGTAAAGAAAGAGATTAACAAGAGTCCCAAGTAAACGC 300 |
| DB | DB | 241 ACACAAATGAGACACAGAGGTGTAAAGAAAGAGATTAACAAGAGTCCCAAGTAAACGC 300 |
| QY | QY | 301 AAAGGGGAGAGACACAGTTTCAAGGGTGTCTATGATCATCTCTTTTTTTTTTTT 360 |
| DB | DB | 301 AAAGGGGAGAGACACAGTTTCAAGGGTGTCTATGATCATCTCTTTTTTTTTTTT 360 |
| QY | QY | 361 TTTTTTTTTTTTTTTTGAACGAGAGTCTGCTCTGTGCCAGGCTGAGAGTCAAGTGGCG 420 |
| DB | DB | 361 TTTTTTTTTTTTTTTTGAACGAGAGTCTGCTCTGTGCCAGGCTGAGAGTCAAGTGGCG 420 |
| QY | QY | 421 GATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTCAAGCCATTTCTTCGCTCAGCCTC 480 |
| DB | DB | 421 GATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTCAAGCCATTTCTTCGCTCAGCCTC 480 |
| QY | QY | 481 CCAATAGACTGGAGCTACAGGCGCCGCCACCTACGCCCCGGCTAATTTTTTTGATTTT 540 |
| DB | DB | 481 CCAATAGACTGGAGCTACAGGCGCCGCCACCTACGCCCCGGCTAATTTTTTTGATTTT 540 |

Db 481 CCAAGTAGCTGGACCTACAGGCGCCGCACTACGCCCCGCTAATTTTTTTTATTTTAA 540
Qy 541 GTAGAGACGGGGTTTCAACCTTTTAAGCCGGGATGGCTTCAATCTCTGACCTCGTAGTCC 600
Db 541 GTAGAGACGGGGTTTCAACCTTTTAAGCCGGGATGGCTTCAATCTCTGACCTCGTAGTCC 600
Qy 601 GCGGCGCTGGGCTCCCAAGTGTGGGATTACAGGCGGTGAGCCAGCGCGCCGGCANTG 660
Db 601 GCGGCGCTGGGCTCCCAAGTGTGGGATTACAGGCGGTGAGCCAGCGCGCCGGCANTG 660
Qy 661 ATCATCTTCTTGAATGCTGATGTGACAGTAAAGCCATCAGACTTACCTTTA 720
Db 661 ATCATCTTCTTGAATGCTGATGTGACAGTAAAGCCATCAGACTTACCTTTA 720
Qy 721 AATATGCACTTTGGGCGCAGGCACTGCTGATGCTGTAATTCAGACCTTTGGAGGC 780
Db 721 AATATGCACTTTGGGCGCAGGCACTGCTGATGCTGTAATTCAGACCTTTGGAGGC 780
Qy 781 AGAGGTGGGTGAATCACTTGAGGCGAGGAGTTTGACACAGCTTGCCCAATGTTGAA 840
Db 781 AGAGGTGGGTGAATCACTTGAGGCGAGGAGTTTGACACAGCTTGCCCAATGTTGAA 840
Qy 841 CTCTGTCTTAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Db 841 CTCTGTCTTAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Qy 901 ACCGTGAATCCAGACTATGCTGAGGCTGAGGCGACAGAGTCACTTGAACCTTGGAGCG 960
Db 901 ACCGTGAATCCAGACTATGCTGAGGCTGAGGCGACAGAGTCACTTGAACCTTGGAGCG 960
Qy 961 GAGGTTCAGTGGGCGGAGATCATATGACGCGCTCCAGGCTGGGCGCAGAGCAAGT 1020
Db 961 GAGGTTCAGTGGGCGGAGATCATATGACGCGCTCCAGGCTGGGCGCAGAGCAAGT 1020
Qy 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Db 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Qy 1081 AAAAAAATATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Db 1081 AAAAAAATATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Qy 1141 AGAATCTAAGGTATATTTGACAAATCATTCAGAACTTTTAAAAAAGAAATCACAGA 1200
Db 1141 AGAATCTAAGGTATATTTGACAAATCATTCAGAACTTTTAAAAAAGAAATCACAGA 1200
Qy 1201 GGCATAGAAAGACAGGAGAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 GGCATAGAAAGACAGGAGAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 AAAAAACAG 1380
Db 1321 AAAAAACAG 1380
Qy 1381 GAGAGAGGGGTGAGAGAGAGAGAGAGATTTGAGAGAGAGCTCAGAAAAGTATGCCAGG 1440
Db 1381 GAGAGAGGGGTGAGAGAGAGAGAGAGATTTGAGAGAGAGCTCAGAAAAGTATGCCAGG 1440
Qy 1441 GAAACACAGAGAGATGGAAGAACTCTGAGAAAAAACAGAGACAAAGATGAAAGAGG 1500
Db 1441 GAAACACAGAGAGATGGAAGAACTCTGAGAAAAAACAGAGACAAAGATGAAAGAGG 1500
Qy 1501 AGTATCGAGGTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAGAGAGAGAGAG 1560
Db 1501 AGTATCGAGGTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAGAGAGAGAGAG 1560
Qy 1561 TCCAGGGGCCAAGATAGTGAACCAAGATTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TCCAGGGGCCAAGATAGTGAACCAAGATTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 1620

Qy 1621 AGGAGGGAGAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCCTCCGGGACAGGGAG 1680
Db 1621 AGGAGGGAGAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCCTCCGGGACAGGGAG 1680
Qy 1681 GTAGAGAGTGAACATTCGGAAGTGGGAGGTGCTCTGGGAGTGAATAGGGAGAGCA 1740
Db 1681 GTAGAGAGTGAACATTCGGAAGTGGGAGGTGCTCTGGGAGTGAATAGGGAGAGCA 1740
Qy 1741 GAGAGAGCTAATTCCTAAGGCGCGATAGGCACTTCAATTCCTCCGGAAATGTCGCCAGGAG 1800
Db 1741 GAGAGAGCTAATTCCTAAGGCGCGATAGGCACTTCAATTCCTCCGGAAATGTCGCCAGGAG 1800
Qy 1801 CAGTGGGTGTTAATCTCAGGCGCCGGTGCAGAGCCCAAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 CAGTGGGTGTTAATCTCAGGCGCCGGTGCAGAGCCCAAGAGAGAGAGAGAGAGAGAG 1860
Qy 1861 AGGACACAGGCTGAGAGAGTGGGCTGAGCTGGAGCAAAATCCCAACCCCTACCTGG 1920
Db 1861 AGGACACAGGCTGAGAGAGTGGGCTGAGCTGGAGCAAAATCCCAACCCCTACCTGG 1920
Qy 1921 GGGACAGGGCAAGTGAAGCTGGTGAAGGTGCTCAGCAGGCAAGGAGAGAGAGTCT 1980
Db 1921 GGGACAGGGCAAGTGAAGCTGGTGAAGGTGCTCAGCAGGCAAGGAGAGAGAGTCT 1980
Qy 1981 GTGCGTCTGACACCAATCTTCTGTCCTCCCTGCTGCTGCTGAGAGGCTGCTAGA 2040
Db 1981 GTGCGTCTGACACCAATCTTCTGTCCTCCCTGCTGCTGCTGAGAGGCTGCTAGA 2040
Qy 2041 CTCTATCTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
Db 2041 CTCTATCTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
Qy 2101 CTGTGGTCTCTCTCTAAGGAGAAATTAAGTGAAGGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 CTGTGGTCTCTCTCTAAGGAGAAATTAAGTGAAGGAGAGAGAGAGAGAGAGAGAG 2160
Qy 2161 GCTCCCGGAGTGGGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGGTGACGGG 2220
Db 2161 GCTCCCGGAGTGGGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGGTGACGGG 2220
Qy 2221 ATGGCTACAG 2280
Db 2221 ATGGCTACAG 2280
Qy 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGAGGTGAGAGAGAGAGAGAG 2340
Db 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGAGGTGAGAGAGAGAGAGAGAG 2340
Qy 2341 CTCTGGGCACTAG 2400
Db 2341 CTCTGGGCACTAG 2400
Qy 2401 GACCGGATGAATGAGTCTCTGTCTGTAATTTGAGTGGTGTGATGATGAGTGGCTGT 2460
Db 2401 GACCGGATGAATGAGTCTCTGTCTGTAATTTGAGTGGTGTGATGATGAGTGGCTGT 2460
Qy 2461 GACTGCCAGGTGTGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
Db 2461 GACTGCCAGGTGTGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
Qy 2521 GACTGCCAGGTGTGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 2521 GACTGCCAGGTGTGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Qy 2581 ATTGT 2640
Db 2581 ATTGT 2640
Qy 2641 GT 2700
Db 2641 GT 2700

| | | | |
|----|------|--|------|
| QY | 2701 | TCCTGAGAGTCCCGGGAATGGGTGCAACAAAGTGTCACTACCAATGGAAGACTGTGACT | 2760 |
| Dp | 2701 | TCCTGAGAGTCCCGGGAATGGGTGCAACAAAGTGTCACTACCAATGGAAGACTGTGACT | 2760 |
| QY | 2761 | GTGTGCTGCTTGCAAGGCGAATTATGTGATTTGTGGCTGAGTGTGAGCGTTATGGAATGCCGTA | 2820 |
| Dp | 2761 | GTGTGCTGCTTGCAAGGCGAATTATGTGATTTGTGGCTGAGTGTGAGCGTTATGGAATGCCGTA | 2820 |
| QY | 2821 | TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGA CTGTGTGTG | 2880 |
| Dp | 2821 | TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGA CTGTGTGTG | 2880 |
| QY | 2881 | TCTGTGTGAGGCCCTGTGTAAATGTCTACTGTATGTGTGATGTGTGCAGCTGTGTCTTGAGT | 2940 |
| Dp | 2881 | TCTGTGTGAGGCCCTGTGTAAATGTCTACTGTATGTGTGATGTGTGCAGCTGTGTGTCTTGAGT | 2940 |
| QY | 2941 | TTCTGTCTCTGCTGAGAGGATATGAGGGGTGACAGGGGTAGCTATCTGTGGAGATGGGTGC | 3000 |
| Dp | 2941 | TTCTGTCTCTGCTGAGAGGATATGAGGGGTGACAGGGGTAGCTATCTGTGGAGATGGGTGC | 3000 |
| QY | 3001 | CAGGTGACTGACTTGCACTGTGTGCTGTGTGTGCAAGAAAGATATGTGGCAGTCTGAACATC | 3060 |
| Dp | 3001 | CAGGTGACTGACTTGCACTGTGTGCTGTGTGTGCAAGAAAGATATGTGGCAGTCTGAACATC | 3060 |
| QY | 3061 | TGTGTCAACACAGGCATCTGTGCGTGTGCACCTAGACACTGTGTGATGAGGGTGTGCACTCC | 3120 |
| Dp | 3061 | TGTGTCAACACAGGCATCTGTGCGTGTGCACCTAGACACTGTGTGATGAGGGTGTGCACTCC | 3120 |
| QY | 3121 | GCTAGGCTGTGCGCGGAGAGTGTGTATACCTGTGAGACAGAGCTGTATGTTAGCTGTGCACCTGTG | 3180 |
| Dp | 3121 | GCTAGGCTGTGCGCGGAGAGTGTGTATACCTGTGAGACAGAGCTGTATGTTAGCTGTGCACCTGTG | 3180 |
| QY | 3181 | GAGGCAACATGTGGCGTGTCTGTGCAGAACCTGTGCGTGTGGCTGTATCTGCTGTTGTGC | 3240 |
| Dp | 3181 | GAGGCAACATGTGGCGTGTCTGTGCAGAACCTGTGCGTGTGGCTGTATCTGCTGTTGTGC | 3240 |
| QY | 3241 | GCGTGTCTTTGGGGTGTGATCTGTGAATGATGTGTGTGCAGGGCCATTCAGCAAGGGTTAA | 3300 |
| Dp | 3241 | GCGTGTCTTTGGGGTGTGATCTGTGAATGATGTGTGTGCAGGGCCATTCAGCAAGGGTTAA | 3300 |
| QY | 3301 | GAAACAGGCGCGGGGCGGTGTGCTACGCTGTAAATCCAGCCCTTTGGAGGCGGAAGCA | 3360 |
| Dp | 3301 | GAAACAGGCGCGGGGCGGTGTGCTACGCTGTAAATCCAGCCCTTTGGAGGCGGAAGCA | 3360 |
| QY | 3361 | GGCGGATTCACCTGAGAGTGTGGAGATTCGAGGCGCACTGTGACCAATGTGAAGAACCCGCTCT | 3420 |
| Dp | 3361 | GGCGGATTCACCTGAGAGTGTGGAGATTCGAGGCGCACTGTGACCAATGTGAAGAACCCGCTCT | 3420 |
| QY | 3421 | CTACTAAATAATACAAAATAATTAGCTGTGTGTGTGTGCGGTGCTGTATATCCAGCTATC | 3480 |
| Dp | 3421 | CTACTAAATAATACAAAATAATTAGCTGTGTGTGTGTGCGGTGCTGTATATCCAGCTATC | 3480 |
| QY | 3481 | GGGAGACTGGGGCGAATAATTCGCTTTAAACCCGGAGAGTGTGAAGTTTCCGTGTGAGCCGAGA | 3540 |
| Dp | 3481 | GGGAGACTGGGGCGAATAATTCGCTTTAAACCCGGAGAGTGTGAAGTTTCCGTGTGAGCCGAGA | 3540 |
| QY | 3541 | TCGAGCAATTGCACTTCAGACTGTGGGCAACAAGACGAAATCTCCGTCTCGAAAGAAAAAA | 3600 |
| Dp | 3541 | TCGAGCAATTGCACTTCAGACTGTGGGCAACAAGACGAAATCTCCGTCTCGAAAGAAAAAA | 3600 |
| QY | 3601 | GAATAAAAAAAGGTTAAGAACCACTGTGAATGGGCAACGGAGGACTGTATGTGAATGGGCG | 3660 |
| Dp | 3601 | GAATAAAAAAAGGTTAAGAACCACTGTGAATGGGCAACGGAGGACTGTATGTGAATGGGCG | 3660 |
| QY | 3661 | ATGCATGTAGTCTGTAGGTCTGTGTGTGTGAGAGAGAGATTGACACGATTTGAGAAAGCAT | 3720 |
| Dp | 3661 | ATGCATGTAGTCTGTAGGTCTGTGTGTGTGAGAGAGAGATTGACACGATTTGAGAAAGCAT | 3720 |
| QY | 3721 | GTTTTCAATCTAGAAATTCAGAAACCTTGAAGGCTGTCCCTCCATAGTGGGCCCTCCCTAG | 3780 |
| Dp | 3721 | GTTTTCAATCTAGAAATTCAGAAACCTTGAAGGCTGTCCCTCCATAGTGGGCCCTCCCTAG | 3780 |
| QY | 3781 | CTGAGCCCTTTCTGTGTGTGTCTGTCTTCCGAAACCTTAGCTCCGCCCATGTAGCTGTAGCC | 3840 |

| | | | |
|----|------|--|------|
| Db | 3781 | CTGAGCCCTTCTTCGTGTCCTCTTCGGAACCTTAGCTCGCCCATGAGCTTGAAC | 3840 |
| OY | 3841 | CCACCTCCTTTCCTCAACCAAGCCCTTAGGCACACTCTAGTGAACCCCGCTAAGGCA | 3900 |
| Db | 3841 | CCACTCTCTTCTCAACCAAGCCCTTAGGCACACTCTAGTGAACCCCGCTAAGGCA | 3900 |
| OY | 3901 | CACCCCTTTGGGCGAGGCTCCACCCCTTATCTGTGGGATCTTCTAGAACCCCTTTCAA | 3960 |
| Db | 3901 | CACCCCTTTGGGCGAGGCTCCACCCCTTATCTGTGGGATCTTCTAGAACCCCTTTCAA | 3960 |
| OY | 3961 | AGTAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGCTGTGTCTCTCCACGGCTGAG | 4020 |
| Db | 3961 | AGTAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGCTGTGTCTCTCCACGGCTGAG | 4020 |
| OY | 4021 | TGCAATGGCGGATCTCGGCTCACTGCAACCTTGCTCCAGGTTCAAGTATTTCTGT | 4080 |
| Db | 4021 | TGCAATGGCGGATCTCGGCTCACTGCAACCTTGCTCCAGGTTCAAGTATTTCTGT | 4080 |
| OY | 4081 | GCTTCACCTCTGAGTAGCTGGGATTAAGGTGCGGCGCACAGCGCTGGCTAATTTTT | 4140 |
| Db | 4081 | GCTTCACCTCTGAGTAGCTGGGATTAAGGTGCGGCGCACAGCGCTGGCTAATTTTT | 4140 |
| OY | 4141 | GTGTCTTTAGTAGAGACAGGGTTTCACTTGTGTGGCGAGGCTGGTCTCAATCCCAAC | 4200 |
| Db | 4141 | GTGTCTTTAGTAGAGACAGGGTTTCACTTGTGTGGCGAGGCTGGTCTCAATCCCAAC | 4200 |
| OY | 4201 | TCAGGTGATCGGCGCACCTCGGCTCCCAAGTCTGGGGTTTCAAGCGTGAAGCACCGC | 4260 |
| Db | 4201 | TCAGGTGATCGGCGCACCTCGGCTCCCAAGTCTGGGGTTTCAAGCGTGAAGCACCGC | 4260 |
| OY | 4261 | CCCCAGCCCAAGTCAGAGCTTTTATATAGAGACTTAACATGTAACTTGAACCTTGACC | 4320 |
| Db | 4261 | CCCCAGCCCAAGTCAGAGCTTTTATATAGAGACTTAACATGTAACTTGAACCTTGACC | 4320 |
| OY | 4321 | CTAATCTAAGTAAATTCCAAAACCCCTTCTGCTCCACCTGAACCCCACTCACTGAAGCC | 4380 |
| Db | 4321 | CTAATCTAAGTAAATTCCAAAACCCCTTCTGCTCCACCTGAACCCCACTCACTGAAGCC | 4380 |
| OY | 4381 | TGACCCCACTCTTGAGAGCAGGTCCTCAATCCCTAAGGCTGTCTCCCTCCCATCCCGAG | 4440 |
| Db | 4381 | TGACCCCACTCTTGAGAGCAGGTCCTCAATCCCTAAGGCTGTCTCCCTCCCATCCCGAG | 4440 |
| OY | 4441 | GCTCCAGCCCCCAAGCTTTGGCACTAACCCCTGAGTGTGTCAAGAACTCTGTAACTCAAT | 4500 |
| Db | 4441 | GCTCCAGCCCCCAAGCTTTGGCACTAACCCCTGAGTGTGTCAAGAACTCTGTAACTCAAT | 4500 |
| OY | 4501 | TTTAACTCTCAGATGTAGTTCTTAGCCAAATTCAGAGAACTGTAGGTCAAGTTAGATCCA | 4560 |
| Db | 4501 | TTTAACTCTCAGATGTAGTTCTTAGCCAAATTCAGAGAACTGTAGGTCAAGTTAGATCCA | 4560 |
| OY | 4561 | GTAACCCCTAGAGCCGTGGGCTGTGTCTTGAACTTGAAGCTTGAAGCTTGAAGATGCA | 4620 |
| Db | 4561 | GTAACCCCTAGAGCCGTGGGCTGTGTCTTGAACTTGAAGCTTGAAGCTTGAAGATGCA | 4620 |
| OY | 4621 | CTCTTATTTCTCAAGGCTGTGCTGTGCCCCCTCAGATGTCAAGAACCCCACTCTAGCT | 4680 |
| Db | 4621 | CTCTTATTTCTCAAGGCTGTGCTGTGCCCCCTCAGATGTCAAGAACCCCACTCTAGCT | 4680 |
| OY | 4681 | GGTCTGGCTCTTAGTGTGAACCAACCCCAAGCCCAAGCCCGCTCTGAAGCCCGGC | 4740 |
| Db | 4681 | GGTCTGGCTCTTAGTGTGAACCAACCCCAAGCCCAAGCCCGCTCTGAAGCCCGGC | 4740 |
| OY | 4741 | CAACCCATTTTCCGTTCCCAAGAGATGTCTCGCAACAATGAATTTTCCCTGAGCAACC | 4800 |
| Db | 4741 | CAACCCATTTTCCGTTCCCAAGAGATGTCTCGCAACAATGAATTTTCCCTGAGCAACC | 4800 |
| OY | 4801 | CCTCTAACAACGCTGCTCTGTGGAGCAACAGACTGGGAGCTGGGGCGGGGAAGAGC | 4860 |
| Db | 4801 | CCTCTAACAACGCTGCTCTGTGGAGCAACAGACTGGGAGCTGGGGCGGGGAAGAGC | 4860 |
| OY | 4861 | CCCGGTGAGTAGACAGACAGCCGCAATCAATGAGATCCGACTGCGATATGCAACC | 4920 |
| Db | 4861 | CCCGGTGAGTAGACAGACAGCCGCAATCAATGAGATCCGACTGCGATATGCAACC | 4920 |

| ID | Sequence | Position | Feature |
|----------|--|--|---------|
| Db | CCCGGTGGATGACGACGACGCGGCATCATCAATGATTCGACCTGCGATATGCAACACC | 4920 | |
| Qy | AGCGGTGGACAGCGCGCGCTGTTGCTAAAGGCCCAACCAAGCTCTACTGCGGGGCGGTGTTGG | 4980 | |
| Db | AGCCCTGGCAGGCCCCGCGCTGTTGCTAAAGGCCCAACCAAGCTCTACTGCGGGGCGGTGTTGG | 4980 | |
| Qy | TGCATCCACAGTGGCGTCTC | 5000 | |
| Db | TGCATCCACAGTGGCTGCTC | 5000 | |
| RESULT 3 | | | |
| ADR72623 | standard; DNA; 11570 BP. | | |
| XX | ADR72623; | | |
| AC | 02-DEC-2004 | (first entry) | |
| XX | | | |
| DT | | | |
| XX | | | |
| DE | Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2. | | |
| XX | | | |
| KX | kallikrein 5; renal cell carcinoma; cytostatic; tumour growth; | | |
| KW | cancer metastasis; chemotherapy; human; serine protease; | | |
| KM | chromosome 19q13.4; KLK5; ds; gene. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 2221..11247 | |
| FT | | /*tag= b | |
| FT | | /product= "Human renal cell carcinoma-related kallikrein | |
| FT | exon | 5 (hK5) protein" | |
| FT | | 2221..2293 | |
| FT | | /*tag= a | |
| FT | intron | /number= 1 | |
| FT | | 2294..4761 | |
| FT | | /*tag= c | |
| FT | exon | /number= 1 | |
| FT | | 4762..5023 | |
| FT | | /*tag= d | |
| FT | | /number= 2 | |
| FT | intron | 5024..5762 | |
| FT | | /*tag= e | |
| FT | | /number= 2 | |
| FT | exon | 5763..6019 | |
| FT | | /*tag= f | |
| FT | | /number= 3 | |
| FT | intron | 6020..6104 | |
| FT | | /*tag= g | |
| FT | exon | /number= 3 | |
| FT | | 6105..6238 | |
| FT | | /*tag= h | |
| FT | | /number= 4 | |
| FT | exon | 6239..11091 | |
| FT | | /*tag= i | |
| FT | | /number= 4 | |
| FT | exon | 11092..11247 | |
| FT | | /*tag= j | |
| FT | | /number= 5 | |
| XX | | | |
| PN | WO2004077060-A2. | | |
| XX | | | |
| PD | 10-SEP-2004. | | |
| XX | | | |
| PF | 26-FEB-2004; 2004WO-CA000280. | | |
| XX | | | |
| PR | 27-FEB-2003; 2003US-0451382P. | | |
| XX | | | |
| PA | (MOUN) MOUNT SINAI HOSPITAL. | | |
| XX | | | |
| PI | Diamonds EP, Petraki CD; | | |
| XX | | | |
| DR | WPI; 2004-662077/64. | | |

| Query Match | 100.0% | Score 5000 | DB 13 | Length 11570 |
|-----------------------|---|--------------|----------|--------------|
| Best Local Similarity | 100.0% | Pred. No. 0 | | |
| Matches 5000 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| DR | P-PSDB; ADR72621. | | | |
| FR | Detecting kallikrein polypeptides/polynucleotides associated with renal | | | |
| PR | cell carcinoma in a patient, for diagnosing/treating the disease, | | | |
| PT | comprises detecting /identifying kallikrein polypeptides/polynucleotides | | | |
| XX | in a sample. | | | |
| PS | Example 1, SEQ ID NO 3; 53bp; English. | | | |
| CC | The invention relates to a novel method for detecting kallikrein | | | |
| CC | polypeptides, or the polynucleotides encoding them, associated with renal | | | |
| CC | cell carcinoma. The method comprises obtaining a sample from a patient | | | |
| CC | and detecting kallikrein polypeptides, or their encoding polynucleotides, | | | |
| CC | where the kallikrein polypeptides are selected from the group consisting | | | |
| CC | of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The | | | |
| CC | detected amounts of the kallikrein polynucleotides are compared to standard | | | |
| CC | amounts. The molecules of the invention demonstrate cytostatic activity | | | |
| CC | whilst the methods and kit may be useful for detecting, characterising, | | | |
| CC | preventing and treating renal cell carcinoma. Furthermore, the methods | | | |
| CC | may be useful for evaluating the probability of the presence of malignant | | | |
| CC | or pre-malignant cells and for detecting and quantifying tumour growth | | | |
| CC | and cancer metastasis. Finally, the methods may be utilised to confirm | | | |
| CC | the absence or removal of all tumour tissue following surgery, cancer | | | |
| CC | chemotherapy and/or radiation therapy and to monitor cancer chemotherapy | | | |
| CC | and tumour reappearance. The current sequence is that of the human | | | |
| CC | ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the | | | |
| CC | invention which encodes a secreted serine protease and is located at | | | |
| CC | chromosome 19q13.4. | | | |
| SO | Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other; | | | |
| OY | GGGCCGAGCGAAGGGAAGAGAGTGAAGCTCCCTGCAAGTGGCTTGATC 60 | | | |
| DB | 1 GGGCCGAGGAGGGAAGGGAAGAGTGAAGCTCCCTGCAAGTGGCTTGATC 60 | | | |
| OY | 61 TCCCTGCGTAAATGCAAGGAGAGGAGCAAGAAACAGAGGAAGAGAGGGGTGGG 120 | | | |
| DB | 61 TCCCTGCGTAAATGCAAGGAGAGGAGCAAGAAACAGAGGAAGAGAGGGGTGGG 120 | | | |
| OY | 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAATTAACAAGAAACAGAGAGAAC 180 | | | |
| DB | 121 AAGAAAGAGAGAGAGAGAGAGAGAGATTAACAATTAACAAGAAACAGAGAGAAC 180 | | | |
| OY | 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAAAGAGAGATAGAGAA 240 | | | |
| DB | 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAAAGAGAGATAGAGAA 240 | | | |
| OY | 241 ACAGAAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGATCCAGATACAC 300 | | | |
| DB | 241 ACAGAAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGATCCAGATACAC 300 | | | |
| OY | 301 AAAGGGGACAGAGCAGATTTCAGGGTGTGTATGATCATCTCTTTT 360 | | | |
| DB | 301 AAAGGGGACAGAGCAGATTTCAGGGTGTGTATGATCATCTCTTTT 360 | | | |
| OY | 361 TTTT 420 | | | |
| DB | 361 TTTT 420 | | | |
| OY | 421 GATTCGGGCTCACTGCAAGCTCCGCTCCGGGTTCACGCAATTCCTGCTTACGCTC 480 | | | |
| DB | 421 GATTCGGGCTCACTGCAAGCTCCGCTCCGGGTTCACGCAATTCCTGCTTACGCTC 480 | | | |
| OY | 481 CCAAGTAGCTGGGACTACAGGCGCCGCACTACGCGCGCTAAATTTTGTATTTT 540 | | | |
| DB | 481 CCAAGTAGCTGGGACTACAGGCGCCGCACTACGCGCGCTAAATTTTGTATTTT 540 | | | |
| OY | 541 GTAGAGCGGGTTTCAACGTTTATGCGGGATGGCTCGATCTCTGACTCGTATTC 600 | | | |
| DB | 541 GTAGAGCGGGTTTCAACGTTTATGCGGGATGGCTCGATCTCTGACTCGTATTC 600 | | | |

D 541 GTAGAGACGGGGTTTCAACCGTTTAAAGCCGGGATGAGCTCGATCTCTGATCC 600
Q 601 GCCCGCTCGGCTCCCAAGTGTCTGGATTAACAGCGTGAAGCCACCGGCCCGGCGCAG 660
D 601 GCCCGCTCGGCTCCCAAGTGTCTGGATTAACAGCGTGAAGCCACCGGCCCGGCGCAG 660
Q 661 ATCATCTTCTGACATACTGATGTGACAACTTAAGCCATCAGACTCTACCTCTTA 720
D 661 ATCATCTTCTGACATACTGATGTGACAACTTAAGCCATCAGACTCTACCTCTTA 720
Q 721 AATATGCAATTTGGGCCAGGCACCGTGCTCATGCTCTTAATTCAGACCTTTGGAGGC 780
D 721 AATATGCAATTTGGGCCAGGCACCGTGCTCATGCTCTTAATTCAGACCTTTGGAGGC 780
Q 781 AGAGGTGGGTGAATCACTTGAGGCCAGAGTTTGAACCGACCTTGCCCAACATGTGAAA 840
D 781 AGAGGTGGGTGAATCACTTGAGGCCAGAGTTTGAACCGACCTTGCCCAACATGTGAAA 840
Q 841 CTCGTCTTTAATAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCTGTGGGGCAC 900
D 841 CTCGTCTTTAATAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCTGTGGGGCAC 900
Q 901 ACCGTGAATCCAGCTATGCTGAGGCTGAGGCAAGAGTCACTTGAACCTTGAGGCG 960
D 901 ACCGTGAATCCAGCTATGCTGAGGCTGAGGCAAGAGTCACTTGAACCTTGAGGCG 960
Q 961 GAGGTTGACGTGGGCCGAGATCAATCACCGCCCTCAGCGTGGGGAGCAGACAGACT 1020
D 961 GAGGTTGACGTGGGCCGAGATCAATCACCGCCCTCAGCGTGGGGAGCAGACAGACT 1020
Q 1021 CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1080
D 1021 CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1080
Q 1081 AAAAAAAAAAATGCTGTCAACAAATGAGCAGAACTGAATTAAGAAATTAATTAATTAAT 1140
D 1081 AAAAAAAAAAATGCTGTCAACAAATGAGCAGAACTGAATTAAGAAATTAATTAATTAAT 1140
Q 1141 AGAATCTAAGGTATTTTGAACAATCATTCAGACCTTTAAAGAAAGAAATTCACAGA 1200
D 1141 AGAATCTAAGGTATTTTGAACAATCATTCAGACCTTTAAAGAAAGAAATTCACAGA 1200
Q 1201 GGCATGAAAACAGGAGAGAAACAGGAGACAGAAAACCTGTGGCCCAAGAGAAACAA 1260
D 1201 GGCATGAAAACAGGAGAGAAACAGGAGACAGAAAACCTGTGGCCCAAGAGAAACAA 1260
Q 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
D 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Q 1321 AAAAAACAG 1380
D 1321 AAAAAACAG 1380
Q 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAAGAAATAGCCGAGGGA 1440
D 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAAGAAATAGCCGAGGGA 1440
Q 1441 GAAACCAAGAGAGATGAGAGAGACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGG 1500
D 1441 GAAACCAAGAGAGATGAGAGAGACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGG 1500
Q 1501 AGTATGAGGGGTGAACAGACAGTGTGGAAATGAGCAAAATGAGCAAGAAAGCAAGCAA 1560
D 1501 AGTATGAGGGGTGAACAGACAGTGTGGAAATGAGCAAAATGAGCAAGAAAGCAAGCAA 1560
Q 1561 TCCAGCGCCCAAGATAAGTGAACCAAGTTGTGAGAAAGCAAGATCTTTAAGGCTGGGG 1620
D 1561 TCCAGCGCCCAAGATAAGTGAACCAAGTTGTGAGAAAGCAAGATCTTTAAGGCTGGGG 1620
Q 1621 AGGCAAGAGAGAGGGGCTGGCTGGCTTCGGAGACCCCTCCCATTTCTCCGGGCGCAGGGAG 1680
D 1621 AGGCAAGAGAGAGGGGCTGGCTGGCTTCGGAGACCCCTCCCATTTCTCCGGGCGCAGGGAG 1680

Q 1681 GTAGGAGATGACATTCGGGACTGGGAGGGGGTGTGCTTGAGGGGTGAGATAGAGGGAGCA 1740
D 1681 GTAGGAGATGACATTCGGGACTGGGAGGGGGTGTGCTTGAGGGGTGAGATAGAGGGAGCA 1740
Q 1741 GAGAGAGCTATTGTCTAAGGCCAGGATAGGCACTTATGGCCGGGAATGTGCCCAAGGAG 1800
D 1741 GAGAGAGCTATTGTCTAAGGCCAGGATAGGCACTTATGGCCGGGAATGTGCCCAAGGAG 1800
Q 1801 CAGTGGGTGTATTAATCTAGGCTCCGGTCCAGAGCCAGAGAGAGAGAGAGAGAGAGAG 1860
D 1801 CAGTGGGTGTATTAATCTAGGCTCCGGTCCAGAGCCAGAGAGAGAGAGAGAGAGAGAG 1860
Q 1861 AGGCAAGGCTGAGAAATGTGCGGCTGAGCTGAGGAGCAAAATCCGCCAATCCCTACTTGG 1920
D 1861 AGGCAAGGCTGAGAAATGTGCGGCTGAGCTGAGGAGCAAAATCCGCCAATCCCTACTTGG 1920
Q 1921 GGGACAGGCAAGTGAAGCTTGTGAGGGTGTCTAGCAGGCAAGGAGAGAGAGAGAGTGTCT 1980
D 1921 GGGACAGGCAAGTGAAGCTTGTGAGGGTGTCTAGCAGGCAAGGAGAGAGAGAGAGTGTCT 1980
Q 1981 GTGCGTCTGCAACCAATCTTTCTGTGCTCCCTCTTGCCCTGTGTGAGGCTGTAGA 2040
D 1981 GTGCGTCTGCAACCAATCTTTCTGTGCTCCCTCTTGCCCTGTGTGAGGCTGTAGA 2040
Q 2041 CTCCTATCTTCAAAATTTCTATGATGCTGAGGCTGAGGCAAGTGCAGATGTGTGCCCGTC 2100
D 2041 CTCCTATCTTCAAAATTTCTATGATGCTGAGGCTGAGGCAAGTGCAGATGTGTGCCCGTC 2100
Q 2101 CTTGTGTCTCTCTCTAAGTGGGAAATAGGTTAGGGAGAGGAGGGGAAATGTGGTTAAG 2160
D 2101 CTTGTGTCTCTCTCTAAGTGGGAAATAGGTTAGGGAGAGGAGGGGAAATGTGGTTAAG 2160
Q 2161 GCTCCCGGATGTGCTGGGCTGCCAACCTCTGACATTCGCCATTCAGGTGCAAGCGGC 2220
D 2161 GCTCCCGGATGTGCTGGGCTGCCAACCTCTGACATTCGCCATTCAGGTGCAAGCGGC 2220
Q 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGAGCTGTGAGCTCTGATCAAGCTTTGCT 2280
D 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGAGCTGTGAGCTCTGATCAAGCTTTGCT 2280
Q 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGGGGTGTGGAGATTGGAGGACTGT 2340
D 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGGGGTGTGGAGATTGGAGGACTGT 2340
Q 2341 CTCTGGGCACTAAGAGCGCTGTCCCCCTGGGGAATGTGTGAGCTGTGGCATGACTCCGG 2400
D 2341 CTCTGGGCACTAAGAGCGCTGTCCCCCTGGGGAATGTGTGAGCTGTGGCATGACTCCGG 2400
Q 2401 GACCGGGTAAATGTAGTCTCTGTCTGTACTTGTGGTGTGGATCGTATGTGGCCCTGT 2460
D 2401 GACCGGGTAAATGTAGTCTCTGTCTGTACTTGTGGTGTGGATCGTATGTGGCCCTGT 2460
Q 2461 GACTGGCAAGTGTGTGTGCTGGGAGGGGATGCTTTTCCATATCAGGTGACTGTGCGG 2520
D 2461 GACTGGCAAGTGTGTGTGCTGGGAGGGGATGCTTTTCCATATCAGGTGACTGTGCGG 2520
Q 2521 CAGGTGGCACTAACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
D 2521 CAGGTGGCACTAACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Q 2581 ATTGT 2640
D 2581 ATTGT 2640
Q 2641 GT 2700
D 2641 GT 2700
Q 2701 TCCCTGAGGTCCCGGATTTGCTGTGCAAAAATGTGTCTATCACTGAAAAGCTGTGACT 2760
D 2701 TCCCTGAGGTCCCGGATTTGCTGTGCAAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760

| | | | |
|----|------|---|------|
| QY | 2761 | GTGTGCTGCTTGCAGGCGCATTAATGTGATTTGTGGCTGAGTGTGACGTTATGGATGCCGTA | 2820 |
| Db | 2761 | GTGTGCTCTTGCAGGCGCATTAATGTGATTTGTGGCTGAGTGTGACGTTATGGATGCCGTA | 2820 |
| QY | 2821 | TTTGTGACCGGTGTGACTGAGAGCTGCTGTAGGGGTGACTGTATGTGACTGTGTGTG | 2880 |
| Db | 2821 | TTTGTGACCGGTGTGACTGAGAGCTGCTGTAGGGGTGACTGTATGTGACTGTGTGTG | 2880 |
| QY | 2881 | TCTGTGTAGGCGCGTGTAAATGCTAACTGATATGTGTGATGTGTGACAGCTGTGTGTGAGT | 2940 |
| Db | 2881 | TCTGTGTAGGCGCGTGTAAATGCTAACTGATATGTGTGATGTGTGACAGCTGTGTGTGAGT | 2940 |
| QY | 2941 | TTCTGTCTCTGTGCTGTGAGGGATAGAGGGTGCAGGGTACTATCTCTGGAGATGGGTGC | 3000 |
| Db | 2941 | TTCTGTCTCTGTGCTGTGAGGGATAGAGGGTGCAGGGTACTATCTCTGGAGATGGGTGC | 3000 |
| QY | 3001 | CAGGTGACTGACTTGTGAGTGTGCTGTGTGTGTGTGTGACAGAAAGTATGTGGAGTCTAAACATC | 3060 |
| Db | 3001 | CAGGTGACTGACTTGTGAGTGTGCTGTGTGTGTGTGTGACAGAAAGTATGTGGAGTCTAAACATC | 3060 |
| QY | 3061 | TGTGTGACACAGGCGCATCTGTGCTGTGTGTGTGTGACAGAAAGTATGTGGAGTCTAAACATC | 3120 |
| Db | 3061 | TGTGTGACACAGGCGCATCTGTGCTGTGTGTGTGTGACAGAAAGTATGTGGAGTCTAAACATC | 3120 |
| QY | 3121 | GCTAGAGCTGCCCGGGAGCGGTGTGTACTGTGAGACAGACCTGTATGTAGTTCAGCTGTG | 3180 |
| Db | 3121 | GCTAGAGCTGCCCGGGAGCGGTGTGTACTGTGAGACAGACCTGTATGTAGTTCAGCTGTG | 3180 |
| QY | 3181 | GAGGCAACATATGGGCGTGTCTGTGACAGAACTGTGCGGTGTGTGTGTGTGTGTGTGTGTG | 3240 |
| Db | 3181 | GAGGCAACATATGGGCGTGTCTGTGACAGAACTGTGCGGTGTGTGTGTGTGTGTGTGTGTG | 3240 |
| QY | 3241 | GCTGTGTCTTGTGGGTGTGATTCGTGTAAATGATGTGTGTGTGTGTGTGTGTGTGTGTGT | 3300 |
| Db | 3241 | GCTGTGTCTTGTGGGTGTGATTCGTGTAAATGATGTGTGTGTGTGTGTGTGTGTGTGTGT | 3300 |
| QY | 3301 | GATCCAGGCGCGGCGCGGT | 3360 |
| Db | 3301 | GATCCAGGCGCGGCGCGGT | 3360 |
| QY | 3361 | GCGCGGATCACTGAGT | 3420 |
| Db | 3361 | GCGCGGATCACTGAGT | 3420 |
| QY | 3421 | CTACTTAAATAATACAAATAATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3480 |
| Db | 3421 | CTACTTAAATAATACAAATAATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3480 |
| QY | 3481 | GAGGAGCTGTGGGCGAGAAATAATCGCTTGAACCGGGAGGTGTGAGGTTCGTGTAGCCGAGA | 3540 |
| Db | 3481 | GAGGAGCTGTGGGCGAGAAATAATCGCTTGAACCGGGAGGTGTGAGGTTCGTGTAGCCGAGA | 3540 |
| QY | 3541 | TTCGGGCGCTTGTGACTTCCAGCTGTGGGCAACAAGCGGAACTTCGCTCTGTGAAAGAAAAA | 3600 |
| Db | 3541 | TTCGGGCGCTTGTGACTTCCAGCTGTGGGCAACAAGCGGAACTTCGCTCTGTGAAAGAAAAA | 3600 |
| QY | 3601 | GAATAAAAAAGGTTAGAACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3660 |
| Db | 3601 | GAATAAAAAAGGTTAGAACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3660 |
| QY | 3661 | ATGCACTGTAGTCTGTAGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3720 |
| Db | 3661 | ATGCACTGTAGTCTGTAGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 3720 |
| QY | 3721 | GTTTTTCATCTGTGAATTCAGAAACCTTAGGCTGTCTTCCCTCGATGTGGCCCCCTTAG | 3780 |
| Db | 3721 | GTTTTTCATCTGTGAATTCAGAAACCTTAGGCTGTCTTCCCTCGATGTGGCCCCCTTAG | 3780 |
| QY | 3781 | CTGAGGCCCTTCTTCTCTGT | 3840 |
| Db | 3781 | CTGAGGCCCTTCTTCTCTGT | 3840 |
| QY | 3841 | GCACCTCTTCTTCTCAACACAGCCCTTAGGCGAGACTGTAGTGACCCCGGCTTAGGCCA | 3900 |

| | | | |
|----|------|---|------|
| Db | 3841 | CCACCTCTTTCCTCAACCAAGCCCTTAAGCCAACTCTAATGAGACCCCGCTTAAGGCA | 3900 |
| Oy | 3901 | CACCCCTTTGGGAGCGCTCACCCCCCTAATCTGTGGGTAACCTTCAAAACCCCTTCAA | 3960 |
| Db | 3901 | CACCCCTTTGGGAGCGGCTCAACCCCTAATCTGTGGGTAACCTTCAAAACCCCTTCAA | 3960 |
| Oy | 3961 | AGTCAGACCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGAG | 4020 |
| Db | 3961 | AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGAG | 4020 |
| Oy | 4021 | TGCAGTGGCGGATCTGGGCTCACTGCAACTCTGCTCCGAGGTTCAAGTATTTCTGT | 4080 |
| Db | 4021 | TGCAGTGGCGGATCTGGGCTCACTGCAACTCTGCTCCGAGGTTCAAGTATTTCTGT | 4080 |
| Oy | 4081 | GCCTTCAACTCTGTAGTGTGGGATTAACAGTGTGGGCAACAGCGCTGGCTAATTTTT | 4140 |
| Db | 4081 | GCCTTCAACTCTGTAGTGTGGGATTAACAGTGTGGGCAACAGCGCTGGCTAATTTTT | 4140 |
| Oy | 4141 | GTTCTTTAGTAGAGACAGGGTTTTCACTTGTGGCAAGCTGTCTCAATCTCCAAAC | 4200 |
| Db | 4141 | GTTCTTTAGTAGAGACAGGGTTTTCACTTGTGGCAAGCTGTCTCAATCTCCAAAC | 4200 |
| Oy | 4201 | TCAGTGTATCGGCCCACTCGGCTCCCAAGTGTGGGGTTACAGCGGTAGACCAACGC | 4260 |
| Db | 4201 | TCAGTGTATCGGCCCACTCGGCTCCCAAGTGTGGGGTTACAGCGGTAGACCAACGC | 4260 |
| Oy | 4261 | CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAAACATGTAAACCTGACCTGGAC | 4320 |
| Db | 4261 | CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAAACATGTAAACCTGACCTGGAC | 4320 |
| Oy | 4321 | CTAACTAAGTCAATTTCCAAACCCCTTCCTGCTCGACGCTGAACCCCACTGAGGAC | 4380 |
| Db | 4321 | CTAACTAAGTCAATTTCCAAACCCCTTCCTGCTCGACGCTGAACCCCACTGAGGAC | 4380 |
| Oy | 4381 | TGACCCCACTTTTGAAGACCAAGTTCATCCTTAAAGCCTGCTCTCCCTCCCATTCGAG | 4440 |
| Db | 4381 | TGACCCCACTTTTGAAGACCAAGTTCATCCTTAAAGCCTGCTCTCCCTCCCATTCGAG | 4440 |
| Oy | 4441 | GCTCAGAGCCCAAGGCTTTGGACATACCCCTGAGCTTGTCAAGAACTCTGTACCCAAAT | 4500 |
| Db | 4441 | GCTCAGAGCCCAAGGCTTTGGACATACCCCTGAGCTTGTCAAGAACTCTGTACCCAAAT | 4500 |
| Oy | 4501 | TTTAACTCTCAATGTAGTCTTAGCCAAATTCAGGAATCTGTAGGTCAGTTAAGTCCA | 4560 |
| Db | 4501 | TTTAACTCTCAATGTAGTCTTAGCCAAATTCAGGAATCTGTAGGTCAGTTAAGTCCA | 4560 |
| Oy | 4561 | GTTAACCTTACTGAGCCTGGGCTCTGTCTTGAAGCTTGAAGCTGGGCTTGAAGGTGCA | 4620 |
| Db | 4561 | GTTAACCTTACTGAGCCTGGGCTCTGTCTTGAAGCTTGAAGCTGGGCTTGAAGGTGCA | 4620 |
| Oy | 4621 | CTTTATTTTTCAGAGGCGCTGCGCCCTGCGCCCTCAGCATGTACACCAACCCCTTAGCT | 4680 |
| Db | 4621 | CTTTATTTTTCAGAGGCGCTGCGCCCTGCGCCCTCAGCATGTACACCAACCCCTTAGCT | 4680 |
| Oy | 4681 | GGTCGGGCTCTTAGTCTGAAACCAACCCCAAGCCCGGCTCTGAGGCCGGCC | 4740 |
| Db | 4681 | GGTCGGGCTCTTAGTCTGAAACCAACCCCAAGCCCGGCTCTGAGGCCGGCC | 4740 |
| Oy | 4741 | CAACCCATTTTCCGTTTCCAGAGCATGTCTTCGCAACAATGATGTTTCTGTGACCAAC | 4800 |
| Db | 4741 | CAACCCATTTTCCGTTTCCAGAGCATGTCTTCGCAACAATGATGTTTCTGTGACCAAC | 4800 |
| Oy | 4801 | CTCTTAAACACGTCGCTTGGGAGCAACAGGACTGGGAGCTGGGGCGGGGAAAGACG | 4860 |
| Db | 4801 | CTCTTAAACACGTCGCTTGGGAGCAACAGGACTGGGAGCTGGGGCGGGGAAAGACG | 4860 |
| Oy | 4861 | CCCGGTGGATGAGCAGCAGCGCATCATATGATCCGACTGGGATATGACACCC | 4920 |
| Db | 4861 | CCCGGTGGATGAGCAGCAGCGCATCATATGATCCGACTGGGATATGACACCC | 4920 |
| Oy | 4921 | AGCCGTGGCAGGCGCGCTGTGCTAAAGCCCAACAGCTTACTGCGGGCGGCTGTTGG | 4980 |

Db 4921 AGCCGTGGCAGCGCGCTGTTCTAAGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
XX 4981 TGCATCCACAGTGGCTGCTC 5000
XX ||||||||||||||||
Db 4981 TGCATCCACAGTGGCTGCTC 5000
RESULT 4
ID ADR72875 standard; DNA; 11570 BP.
XX ADR72875;
XX ADR72875;
XX 02-DEC-2004 (first entry)
DE Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.
XX kallikrein 5; tumour marker; ovarian cancer;
KW epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
KW KLK5; ds; gene.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 221..11247
XX FT /tag= b
XX FT /product= "Human ovarian cancer-related tumour marker
XX FT kallikrein 5 (hK5) protein"
XX FT 221..2293
XX FT /tag= a
XX FT /number= 1
XX FT 2294..4761
XX FT /tag= c
XX FT /number= 1
XX FT 4762..5023
XX FT /tag= d
XX FT /number= 2
XX FT 5024..5762
XX FT /tag= e
XX FT /number= 2
XX FT 5763..6019
XX FT /tag= f
XX FT /number= 3
XX FT 6020..6104
XX FT /tag= g
XX FT /number= 3
XX FT 6105..6238
XX FT /tag= h
XX FT /number= 4
XX FT 6239..11091
XX FT /tag= i
XX FT /number= 4
XX FT 11092..11247
XX FT /tag= j
XX FT /number= 5
XX
XX WO2004075713-A2.
XX
XX 10-SEP-2004.
XX
XX 26-FEB-2004; 2004WO-CA000281.
XX
XX 26-FEB-2003; 2003US-0450406P.
XX
XX (MOUN) MOUNT SINAI HOSPITAL.
XX
XX Diamandis BP;
XX
XX WPI; 2004-661815/64.
XX
XX P-PSDB; ADR72873.
XX
XX Kallikrein markers detection method for detecting ovarian cancer in
XX patient, involves detecting kallikrein markers and CA125 in sample
PT

PT obtained from patient, and comparing detected amounts with standard
XX amounts.
XX
XX Example 2; SEQ ID NO 5; 102pp; English.
XX
XX The invention relates to a novel method for detecting a plurality of
XX kallikrein markers associated with ovarian cancer. The method comprises
XX obtaining a sample from a patient and detecting in the sample a plurality
XX of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
XX the kallikrein markers are selected from the group consisting of
XX kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
XX kallikrein 11. The detected amounts of the kallikrein markers are
XX compared with standard amounts. The method of the invention may be useful
XX for detecting kallikrein markers associated with ovarian cancer in a
XX patient and thus for detecting ovarian cancer, particularly epithelial
XX ovarian carcinoma. The current sequence is that of the human ovarian
XX cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention
XX which encodes a secreted serine protease and is located at chromosome
XX 19q13.4.
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 5000; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCAGAGTGAAGGCAAGAGAGAGTGGAGCTCCCTGCAAGTGGCTTGAGTC 60
Db 1 GGGCCAGAGTGAAGGCAAGAGAGAGTGGAGCTCCCTGCAAGTGGCTTGAGTC 60
QY 61 TCCCTGCTTAATATGACGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTGCTTAATATGACGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 AAGAAAG 180
Db 121 AAGAAAG 180
QY 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAG 240
QY 241 ACACAAATGAGACACAGAGGTGTAAGAAAGAGATTAACAGATCCACAGTACAG 300
Db 241 ACACAAATGAGACACAGAGGTGTAAGAAAGAGATTAACAGATCCACAGTACAG 300
QY 301 AAAGGGGAGAGACAGATTTTCAAGGTGTCTATGATCATCTTTTCTTTTCTTT 360
Db 301 AAAGGGGAGAGACAGATTTTCAAGGTGTGTCTATGATCATCTTTTCTTTTCTTT 360
QY 361 TTTTCTTTTCTTTTGAAGCGAGTCTGCTGTCTGCGCCAGGCTGAGTGGAGCGG 420
Db 361 TTTTCTTTTCTTTTGAAGCGAGTCTGCTGTCTGCGCCAGGCTGAGTGGAGCGG 420
QY 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGCGGTTCAAGCCATTCTCTGACGCTC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGCGGTTCAAGCCATTCTCTGACGCTC 480
QY 481 CCAAGTAGCTGGAGCTACAGCGCCGCCCACTAGCGCCGCGCTAATTTTGTATTTT 540
Db 481 CCAAGTAGCTGGAGCTACAGCGCGCCGCCCACTAGCGCCGCGCTAATTTTGTATTTT 540
QY 541 GTAGAGAGGGGTTTCAACGTTTGAAGCGGAGTGGCTGAGTCTGAGTCTGAGTCC 600
Db 541 GTAGAGAGGGGTTTCAACGTTTGAAGCGGAGTGGCTGAGTCTGAGTCTGAGTCC 600
QY 601 GCCCGCTCGGCTCCCAAAGTCTGGAGTTACAGCGCTGAGCCACCGCGCCGCAATG 660
Db 601 GCCCGCTCGGCTCCCAAAGTCTGGAGTTACAGCGCTGAGCCACCGCGCCGCAATG 660
QY 661 ATCATCTTCTTGACTATGCTGATGTGACAGTACTTAAGCCATGAGCTTACCTTTTA 720
Db 661 ATCATCTTCTTGACTATGCTGATGTGACAGTACTTAAGCCATGAGCTTACCTTTTA 720

QY 721 AATATGAGTTGGGCCAGGCAACGGTGTCTCATGCTGTAAATTCACAGACTTTGGGAGGC 780
DB 721 AATATGAGTTGGGCCAGGCAACGGTGTCTCATGCTGTAAATTCACAGACTTTGGGAGGC 780
QY 781 AGAGTGGGTGAATCATTGAGGCCAGAGTTTGAGACAGCTGGCCAAATGTTGAAA 840
DB 781 AGAGTGGGTGAATCATTGAGGCCAGAGTTTGAGACAGCTGGCCAAATGTTGAAA 840
QY 841 CTGTCTTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
DB 841 CTGTCTTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
QY 901 ACCGTGATCCGAGCTATGCTGAGAGGCTGAGGCAAGAGTCACTTGAACCTTGAGAGCG 960
DB 901 ACCGTGATCCGAGCTATGCTGAGAGGCTGAGGCAAGAGTCACTTGAACCTTGAGAGCG 960
QY 961 GAGGTGACAGTGGGCCAGATCACATCACCGCCCTCCAGCCTGGGCCAGAGCAAGACT 1020
DB 961 GAGGTGACAGTGGGCCAGATCACATCACCGCCCTCCAGCCTGGGCCAGAGCAAGACT 1020
QY 1021 CTGTCTCAAAATTAATTAATTAACCAACGAAACAGTGTGTGTACCTTAGTTATATCT 1080
DB 1021 CTGTCTCAAAATTAATTAATTAACCAACGAAACAGTGTGTGTACCTTAGTTATATCT 1080
QY 1081 AAAAAAATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
DB 1081 AAAAAAATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
QY 1141 AGAATCTAAGGTATATTTTGAACAATCATTCAGAACCTTTAAAAAAGAAATCAGAGA 1200
DB 1141 AGAATCTAAGGTATATTTTGAACAATCATTCAGAACCTTTAAAAAAGAAATCAGAGA 1200
QY 1201 GGCATAGAAAAGACAGGAGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
DB 1201 GGCATAGAAAAGACAGGAGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
QY 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGTGAATGAGACAGACAGAG 1320
DB 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGTGAATGAGACAGACAGAG 1320
QY 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGACAGAGAGAGAGAGAGTGAAGAA 1380
DB 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGACAGAGAGAGAGAGAGTGAAGAA 1380
QY 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAGAAAATATGCGAGGGA 1440
DB 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAGAAAATATGCGAGGGA 1440
QY 1441 GAACCAACAGAGAGATTCAGAAAGACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGG 1500
DB 1441 GAACCAACAGAGAGATTCAGAAAGACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGG 1500
QY 1501 AGTATCGAGGGGTGAACAGACAGTGTGGAATGAGCAAAATGCAGAGAAAGAAAGCAAG 1560
DB 1501 AGTATCGAGGGGTGAACAGACAGTGTGGAATGAGCAAAATGCAGAGAAAGAAAGCAAG 1560
QY 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTGAGAAAGCCAGATCCTTAAAGGCTGGGG 1620
DB 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTGAGAAAGCCAGATCCTTAAAGGCTGGGG 1620
QY 1621 AGGCAAGGAGAGGGGTGGCTGGCTTCGAGAGACCCCTCCCATTTCTCCGGGCCAGAGAG 1680
DB 1621 AGGCAAGGAGAGGGGTGGCTGGCTTCGAGAGACCCCTCCCATTTCTCCGGGCCAGAGAG 1680
QY 1681 GTAGGAGATGACATTTCCGAGCTGGGTGGGGGTGTCTCTGGGGGTGAGAGTAGGGGAGCA 1740
DB 1681 GTAGGAGATGACATTTCCGAGCTGGGTGGGGGTGTCTCTGGGGGTGAGAGTAGGGGAGCA 1740
QY 1741 GGAGAGCTATGTGCTAAAGCCCGATAGGCACTCATTTGCCGGGAAATGTGCCAGAGAG 1800
DB 1741 GGAGAGCTATGTGCTAAAGCCCGATAGGCACTCATTTGCCGGGAAATGTGCCAGAGAG 1800

QY 1801 CAGTGGGTGTTATTAATCAGAGCCCGTGGCCAGAGCCCAAGAGAGAGCAATGGCCAGGA 1860
DB 1801 CAGTGGGTGTTATTAATCAGAGCCCGTGGCCAGAGCCCAAGAGAGAGCAATGGCCAGGA 1860
QY 1861 AGGCAACAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCAACCTTCACTGG 1920
DB 1861 AGGCAACAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCAACCTTCACTGG 1920
QY 1921 GGGACAGGAGCAAGTGAACCTGTGAGAGGGTGTGACAGAGAGAGAGAGAGAGTGTCT 1980
DB 1921 GGGACAGGAGCAAGTGAACCTGTGAGAGGGTGTGACAGAGAGAGAGAGAGAGTGTCT 1980
QY 1981 GTGGGTCTGACACCAATCTTTCTGTCCCTCTTGGCCCTGTGTGAGAGCTGTAGA 2040
DB 1981 GTGGGTCTGACACCAATCTTTCTGTCCCTCTTGGCCCTGTGTGAGAGCTGTAGA 2040
QY 2041 CTCCTATCTTCTGAATTTATTAAGTCTGGGTCTCAGCGAGTGCAGATGTGGCCCTG 2100
DB 2041 CTCCTATCTTCTGAATTTATTAAGTCTGGGTCTCAGCGAGTGCAGATGTGGCCCTG 2100
QY 2101 CTGTGTGTTCTCTCTACCTGGGAAATTAAGTAAAGGAGAGAGAGAGAGTGAAG 2160
DB 2101 CTGTGTGTTCTCTCTACCTGGGAAATTAAGTAAAGGAGAGAGAGAGAGTGAAG 2160
QY 2161 GCTCCCGGATGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGACAGCGGCC 2220
DB 2161 GCTCCCGGATGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGACAGCGGCC 2220
QY 2221 ATGGCTACAGCAAGACCCCTGGAATGAGGTCTGTGCTGTGATCAGACCTTGCTT 2280
DB 2221 ATGGCTACAGCAAGACCCCTGGAATGAGGTCTGTGCTGTGATCAGACCTTGCTT 2280
QY 2281 CTGGGGGTGACAGGTAAACAGAACTGTGGGTGGAGAGGTTTGGAGTGAAGTGT 2340
DB 2281 CTGGGGGTGACAGGTAAACAGAACTGTGGGTGGAGAGGTTTGGAGTGAAGTGT 2340
QY 2341 CTGTGCGGCACTAGAGGCTGTCCCTGGGAAATGTGTGAGAGCTGGGATGACTCCGG 2400
DB 2341 CTGTGCGGCACTAGAGGCTGTCCCTGGGAAATGTGTGAGAGCTGGGATGACTCCGG 2400
QY 2401 GACCGGGTGAATGATGATCTGTCTGTACTTGTGTGTGCAATGCAATGTGGCCCTGT 2460
DB 2401 GACCGGGTGAATGATGATCTGTCTGTACTTGTGTGTGCAATGCAATGTGGCCCTGT 2460
QY 2461 GACTGCCACGGTGTGTGTGGGAGAGGGAGATGCTTTCCCATATACAGGTGACTGTGCGG 2520
DB 2461 GACTGCCACGGTGTGTGTGGGAGAGGGAGATGCTTTCCCATATACAGGTGACTGTGCGG 2520
QY 2521 CAGGTGCACTGACCCCTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
DB 2521 CAGGTGCACTGACCCCTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 2581 ATTGTGTGTGATCTCCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
DB 2581 ATTGTGTGTGATCTCCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
QY 2641 GTGTGTGCTGT 2700
DB 2641 GTGTGTGCTGT 2700
QY 2701 TCCCTGAGGTCCCGGAGTTGCTGTGCAACAAAGGTGATCACCATGAGAAAGCTGTGACT 2760
DB 2701 TCCCTGAGGTCCCGGAGTTGCTGTGCAACAAAGGTGATCACCATGAGAAAGCTGTGACT 2760
QY 2761 GTGTGCTGTGTGCAAGGATTAATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGT 2820
DB 2761 GTGTGCTGTGTGCAAGGATTAATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGT 2820
QY 2821 TTTGTGACCGTGTGACATCTGAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
DB 2821 TTTGTGACCGTGTGACATCTGAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
QY 2881 TCTGTGTGAGGCGGTGAATGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940

2881 TCTGTGTGAGGCGCGTGAATGCTACTGTATGTGTGTCAGAGCTGTGTCTGAGAT 2940
2941 TTCTGTCTGCGCTGAGGGATAGAGGGTGCAGAGGGTACTATCTCTGGAGATGTGGTC 3000
2941 TTCTGTCTGCGCTGAGGGATAGAGGGTGCAGAGGGTACTATCTCTGGAGATGTGGTC 3000
3001 CAGGTGACTGACTTTGAGAGTGTGTGCTGTGTGCAAGAGATATGTGGCACTGAACTATC 3060
3001 CAGGTGACTGACTTTGAGAGTGTGTGCTGTGTGCAAGAGATATGTGGCACTGAACTATC 3060
3061 TGTGCACACAGGCACTGTGTGCTGTGCACTGAGACACTGTGTGATGAGGGTGTCCATCCC 3120
3061 TGTGCACACAGGCACTGTGTGCTGTGCACTGAGACACTGTGTGATGAGGGTGTCCATCCC 3120
3121 GCTAGGCTGCGCGGAGCGGTGTGACCTGTGAGACAGAGCTGTATGTAGCTGACCTGTG 3180
3121 GCTAGGCTGCGCGGAGCGGTGTGACCTGTGAGACAGAGCTGTATGTAGCTGACCTGTG 3180
3181 GAGGCAACATGAGCGCTGTCTGACAGAACTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTG 3240
3181 GAGGCAACATGAGCGCTGTCTGACAGAACTGCGTGTGCTGTGCTGTGCTGTGCTGTGCTG 3240
3241 GCGTGTCTTGTGGGCTGAGTGTGTGTATGATGATGATGATGATGATGATGATGATGATG 3300
3241 GCGTGTCTTGTGGGCTGAGTGTGTGTATGATGATGATGATGATGATGATGATGATGATG 3300
3301 GAAACAGCGCGGCGGCGGTGTGCTGACGCGCTGTATATCCAGCGCTTGTGGGCGCGAGGA 3360
3301 GAAACAGCGCGGCGGCGGTGTGCTGACGCGCTGTATATCCAGCGCTTGTGGGCGCGAGGA 3360
3361 GGGGATCACTGAGGTGCGGAGATCGAGGCGAGCGCTGACCAACATGAGAAACCCGCTGT 3420
3361 GGGGATCACTGAGGTGCGGAGATCGAGGCGAGCGCTGACCAACATGAGAAACCCGCTGT 3420
3421 CTACTTAAATAATCAAAAAATTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
3421 CTACTTAAATAATCAAAAAATTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
3421 CTACTTAAATAATCAAAAAATTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3480
3481 GGGAGACTGGGGCAGAAAAATCGCTTGAACCGGAGAGGTGAGAGGTGTGGGCTGAGCGGAG 3540
3481 GGGAGACTGGGGCAGAAAAATCGCTTGAACCGGAGAGGTGAGAGGTGTGGGCTGAGCGGAG 3540
3481 GGGAGACTGGGGCAGAAAAATCGCTTGAACCGGAGAGGTGAGAGGTGTGGGCTGAGCGGAG 3540
3541 TCGCGCATTTGCACTCGAGCTGTGGGCAACAGAGCGAACTCGCTGTGAAAGAAAAA 3600
3541 TCGCGCATTTGCACTCGAGCTGTGGGCAACAGAGCGAACTCGCTGTGAAAGAAAAA 3600
3601 GAAAAAAGAGGTGAAGAACAGTGAATGGGACCGGAGAGCTGATGATGATGATGATGATG 3660
3601 GAAAAAAGAGGTGAAGAACAGTGAATGGGACCGGAGAGCTGATGATGATGATGATGATG 3660
3601 GAAAAAAGAGGTGAAGAACAGTGAATGGGACCGGAGAGCTGATGATGATGATGATGATG 3660
3661 ATGCATGT 3720
3661 ATGCATGT 3720
3721 GTTTTCACTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGTGTGTGTGTGTGT 3780
3721 GTTTTCACTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGTGTGTGTGTGTGT 3780
3781 CTGAGCGCTTTCTTCTGCTGT 3840
3781 CTGAGCGCTTTCTTCTGCTGT 3840
3841 CCACCTCTTTCTCTCAACACGCGCTAGGCGAGACTGTAGTGAACCGCGCTTAAGGCGCA 3900
3841 CCACCTCTTTCTCTCAACACGCGCTAGGCGAGACTGTAGTGAACCGCGCTTAAGGCGCA 3900
3901 CACCCCTTTGTGGGCAAGGCTCAACCCCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
3901 CACCCCTTTGTGGGCAAGGCTCAACCCCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
3961 AGTCAGAGCTTTTGT 4020

3961 AGTCAGAGCTTTTGT 4020
4021 TGCAGTGGGCTATCTCGGCTCAGTCAACCTGTGCTCCAGGTTCAAGTATTTCTGT 4080
4021 TGCAGTGGGCTATCTCGGCTCAGTCAACCTGTGCTCCAGGTTCAAGTATTTCTGT 4080
4081 GCGTCACTCTGTAGTACGTGGGATTAACAGGTGTGCGGCAACCGCTGTGCTAATTTT 4140
4081 GCGTCACTCTGTAGTACGTGGGATTAACAGGTGTGCGGCAACCGCTGTGCTAATTTT 4140
4081 GCGTCACTCTGTAGTACGTGGGATTAACAGGTGTGCGGCAACCGCTGTGCTAATTTT 4140
4141 GTGTCTTTGT 4200
4141 GTGTCTTTGT 4200
4141 GTGTCTTTGT 4200
4201 TCAGGTGATCGGCGCACTGTGCGCTCCAGAGGTGTGCGGTTTACAGGCGTGAAGCGC 4260
4201 TCAGGTGATCGGCGCACTGTGCGCTCCAGAGGTGTGCGGTTTACAGGCGTGAAGCGC 4260
4201 TCAGGTGATCGGCGCACTGTGCGCTCCAGAGGTGTGCGGTTTACAGGCGTGAAGCGC 4260
4261 CCCAGCCCAAGTCAAGACTCTTTATAGAGACTCTTAACATGTAACTCTGACCCCTGGC 4320
4261 CCCAGCCCAAGTCAAGACTCTTTATAGAGACTCTTAACATGTAACTCTGACCCCTGGC 4320
4321 CTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCGCTTGAACCCCACTCACTGAGG 4380
4321 CTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCGCTTGAACCCCACTCACTGAGG 4380
4381 TGACCCCACTTTCTGTAGACAGTTCATCCCTAAAGCCCTGTGCTCCCTCCATCCGAG 4440
4381 TGACCCCACTTTCTGTAGACAGTTCATCCCTAAAGCCCTGTGCTCCCTCCATCCGAG 4440
4441 GCTCAGCCGCCCAAGCTTTGTGACACTACCCCTGAGCTTGTCCAGAAATCTGTACCAAT 4500
4441 GCTCAGCCGCCCAAGCTTTGTGACACTACCCCTGAGCTTGTCCAGAAATCTGTACCAAT 4500
4501 TTTTACCTCAATGTAGT 4560
4501 TTTTACCTCAATGTAGT 4560
4561 GTTACCCCTACCTGAGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4620
4561 GTTACCCCTACCTGAGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4620
4621 CTCTTATTTCTCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
4621 CTCTTATTTCTCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
4681 GGTCTGGCTCTTGTGCTGTGAACCCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740
4681 GGTCTGGCTCTTGTGCTGTGAACCCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740
4741 CAACCATTTTCTGTTCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4800
4741 CAACCATTTTCTGTTCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4800
4801 CCTTAAACCGTGTGCTCTGTGAGCAACAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 4860
4801 CCTTAAACCGTGTGCTCTGTGAGCAACAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 4860
4801 CCTTAAACCGTGTGCTCTGTGAGCAACAGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 4860
4861 CCCGCTGAGTACACAGCAGAGCGGCTATCATATGATGATGATGATGATGATGATGATGATG 4920
4861 CCCGCTGAGTACACAGCAGAGCGGCTATCATATGATGATGATGATGATGATGATGATGATG 4920
4921 AGCGGTGAGGCGGCGGCTGT 4980
4921 AGCGGTGAGGCGGCGGCTGT 4980
4981 TGCATCAACAGTGTGCTGCT 5000

RESULT 5
ACH80345

ID ACH80345 standard; DNA; 525 BP.
AC ACH80345;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #13540.
XX
KW Human; probe; seq; gene expression; single exon probe; microarray;
KM alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
PN US2003194704-A1.
XX
PD 16-OCT-2003.
PF 03-APR-2002; 2002US-00029386.
PR 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
PT Penn SG, Rank DR, Hanzel DK;
XX
XX WPI: 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
XX surveying tissues.
XX
XX
PS Claim 15; SEQ ID NO 13540; 80bp; English.
XX
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above, a
CC method of sealing and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subcription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, method and apparatus are useful in gene
CC expression analysis. The probe may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
XX Sequence 525 BP; 99 A; 178 C; 145 G; 103 T; 0 U; 0 Other;

| Query Match | Similarity | 9.4%; | Score 472; | DB 12; | Length 525; |
|--------------|--------------|--|--------------------|--------|-------------|
| Best Local | Similarity | 100.0%; | Pred. No. 4.3e-76; | | |
| Matches 472; | Conservative | 0; | Mismatches | 0; | Indels |
| | | | | | Gaps |
| | | | | | 0; |
| Qy | 4529 | TCCAGGATCTGTGAGGTCCAGTTAGAGTCCAGTAACTTACCTGAGCTGAGCTCTGTC | 4588 | | |
| Db | 1 | TCCAGGATCTGTGAGGTCCAGTTAGAGTCCAGTAACTTACCTGAGCTGAGCTCTGTC | 60 | | |
| Qy | 4589 | CTTGAAGCTTGAAGCTTGAAGGTGCCACTTTATTCTCGAGGCTTGCCTCTGTC | 4648 | | |
| Db | 61 | CTTGAAGCTTGAAGCTTGAAGGTGCCACTTTATTCTCGAGGCTTGCCTCTGTC | 120 | | |
| Qy | 4649 | CCCTCAGATGTGAGACCCACCCTCTAGCTGCTGGCTCTTGAAGTGAAGAACAC | 4708 | | |
| Db | 121 | CCCTCAGATGTGAGACCCACCCTCTAGCTGCTGGCTCTTGAAGTGAAGAACAC | 180 | | |
| Qy | 4709 | CCCCAGGCCAAGCCCGCTCTGAGCCCGCCCAACCATTTTCCGTTCCAGACATGT | 4768 | | |
| Db | 181 | CCCCAGGCCAAGCCCGCTCTGAGCCCGCCCAACCATTTTCCGTTCCAGACATGT | 240 | | |
| Qy | 4769 | TCTTGCCCAATATATGTTTCTGTGACACCCCTCTTAAACCGTGGCTCTTGGAGCA | 4828 | | |
| Db | 241 | TCTTGCCCAATATATGTTTCTGTGACACCCCTCTTAAACCGTGGCTCTTGGAGCA | 300 | | |
| Qy | 4829 | CCAGGACCTGGAGCTGGGGCCGGGGAAGAAGCCCGGTGCGATATGACAGACAGCCGAT | 4888 | | |
| Db | 301 | CCAGGACCTGGAGCTGGGGCCGGGGAAGAAGCCCGGTGCGATATGACAGACAGCCGAT | 360 | | |
| Qy | 4889 | CATCAATGATTCGCACTGCGATATATGACACCCAGCCGTGACAGGCGCGCTGTTGTAAG | 4948 | | |
| Db | 361 | CATCAATGATTCGCACTGCGATATATGACACCCAGCCGTGACAGGCGCGCTGTTGTAAG | 420 | | |
| Qy | 4949 | GCCCAACCAAGCTCTACTGCGGGGCGGTGTTGATGATCCACAGTGGCTGCTC | 5000 | | |
| Db | 421 | GCCCAACCAAGCTCTACTGCGGGGCGGTGTTGATGATCCACAGTGGCTGCTC | 472 | | |
| RESULT 6 | | | | | |
| ACN44942 | | | | | |
| ID | ACN44942 | standard; DNA; 41454 BP. | | | |
| XX | AC | | | | |
| XX | AC | ACN44942; | | | |
| XX | DT | 18-NOV-2004 (first entry) | | | |
| DE | DE | Human genomic sequence hCG21040. | | | |
| XX | XX | Cytosstatic; carcinoma; lymphoma; cancer; human; gene; sb. | | | |
| KW | XX | Homo sapiens. | | | |
| XX | OS | | | | |
| XX | PN | WO2003073826-A2. | | | |
| XX | DD | 12-SEP-2003. | | | |
| XX | PE | 28-FEB-2003; 2003WO-US006235. | | | |
| XX | PR | 01-MAR-2002; 2002US-00087192. | | | |
| PA | XX | (SAGR-) SAGRES DISCOVERY. | | | |
| XX | XX | Morris DW; | | | |
| XX | PI | | | | |
| XX | DR | WPI; 2003-328604/31. | | | |
| XX | XX | | | | |
| XX | PT | Recombinant nucleic acid useful for diagnosis and treatment of carcinoma | | | |
| XX | XX | comprises a nucleotide sequence. | | | |
| XX | XX | Claim 1; SEQ ID NO 1642; Opp; English. | | | |
| XX | XX | The present invention relates to novel DNA and protein sequences which | | | |
| CC | CC | are associated with carcinomas. The sequences are useful for: (1) for | | | |

| | | | |
|---|-------|---|-------|
| OY | 314 | CACGTTTTCAGGGTGGGTCATGATCATCTCTTTTTTTTTTTTTTTTTTTTTTTT | 373 |
| Db | 20797 | CACAGTGCCTATTTTCCCTTCCAGTATGATTCCTTTTTTTTTTTTTTTTTTTTTT | 20738 |
| OY | 374 | TTGAGACGAGATCTCGCTCTGTGCGCCACGAGCTGAGTGCAGTGGAGATCTCGCTCAC | 433 |
| Db | 20737 | TTGAGACGAGATCTCGCTCTGTGCGCCACGAGCTGAGTGCAGTGGAGATCTTGGCTCAC | 20678 |
| OY | 434 | TGCAAGCTCCGCTCTCCGGGTTACAGGCAATCTCTGCTCAGCTTCCCAAGTACTGGG | 493 |
| Db | 20677 | TGCAAGCTCCGCTCTCCGGGTTACAGGCAATCTCTGCTCAGCTTCCCAAGTACTGGG | 20618 |
| OY | 494 | ACTACAGGCGCCGCGCACTACGCGCCGAGTAAATTTTTTGTATTTTATGATGAGACGGAGT | 553 |
| Db | 20617 | ACGACAGGCGCCGCGCACTACGCGCCGAGTAAATTTTTTGTATTTTATGATGAGACGGAGT | 20559 |
| OY | 554 | TTCAACGTTTTTACCGGAGATGGCTCGATCTCTGACCTTGATACCGCCCGCTCGGCC | 613 |
| Db | 20558 | TTCAACGTTTTTACCGGAGATGGCTCGATCTCTGACCTTGATACCGCCCGCTCGGCC | 20499 |
| OY | 614 | TCCCAAAAGTGTGGGATTTACAGGCTGTAGCCACCGCGCCCGGCAATCATCT----- | 667 |
| Db | 20498 | TCCCAAAAGTGTGGGATTTACAGGCTGTAGCCACCGCGCCCGGCAATCATCTTTCTTAA | 20439 |
| OY | 668 | -----TCTTGACTAGCTGATGTGCAAGTAC | 694 |
| Db | 20438 | TGTGCTTATTTTGGTTTAACTTCTTATTTTGAAGTTTTTGGATATACATGAAACAAGCAC | 20379 |
| OY | 695 | CTTAAAGCATCA-GACTCTACCTTTTAAATATGACAGTTTGGGCGACGACCGTGGCTCAT | 753 |
| Db | 20378 | AAATTAATATCATCTATCTCAACCGTCTTAAATATGATTTTGAAGCTCAAGGACCGGTGACTCAC | 20319 |
| OY | 754 | GCTGTATTTTCAGACATTTTGGGAGGCGAGAGTGGGTGATCATCTTGAAGCCAGAGATT | 813 |
| Db | 20318 | ACCTGTATCTCCCAACACTTTTGAAGAGGCGGAGTGGGCACTCAC--GAGGTCAAGAGTTTC | 20261 |
| OY | 814 | GAGACCAAGCTGGGCAACATGTGTAATCTGTCTTTACTTAAAAAATTTTTTTTTTTT | 873 |
| Db | 20260 | GAGACCAAGCTGGGCAACATGTGTAATCTGTCTTTACTTAAAAAATTTTTTTTTTTT | 20213 |
| OY | 874 | AAAAATCAGCCGGGTGTGTTGGGCGACACCTGTATCTCCAGCTATCTGAGAGCTGAGGC | 933 |
| Db | 20212 | AAAAATCAGCTGGGTGTGTTGGGCGACCGCTGTATCTCCAGCTATCTGAGAGCTGAGGC | 20153 |
| OY | 934 | ACGAGATCTCACTTGAACCTCTGAGGCGGAGTTGCACTGGGCGGAGATCATCACCGCC | 993 |
| Db | 20152 | ATGGAATTTCTTGAAGCTCCGGGAGGCGAGAGTTGCACTGAGGAGACATCCATTGTA | 20093 |
| OY | 994 | CTTCAGGCTGGGCGACAGACAAAGACTGTCTCTCAATATAATTAATTAACAAACGACAA | 1053 |
| Db | 20092 | GTCACGCTGGGCGACAGACGAGACTCTCTCTCAATATAATTAATTAATTAATTAAGTT | 20033 |
| OY | 1054 | GCAGTTTGTGTACCTTAGTTATATCTAAAAAATAAT | 1091 |
| Db | 20032 | GTTGATCTTTCACCAATCATTTAATGATGATTAAGAT | 19995 |
| RESULT 8 | | | |
| ADN94799 | | | |
| ID ADN94799 standard; DNA; 91352 BP. | | | |
| AC ADN94799; | | | |
| AT | | | |
| TT 12-AUG-2004 (first entry) | | | |
| TT | | | |
| DE DNA encoding human nidogen. | | | |
| XX | | | |
| KW gene; human; Antisense Therapy; Gene Therapy; nidogen; | | | |
| KW Chedick-Higashi syndrome; de. | | | |
| XX | | | |
| SS Homo sapiens | | | |
| XX | | | |

| | |
|---------------------------|--|
| PN | US2004097451-A1. |
| PX | |
| PD | 20-MAY-2004. |
| PX | |
| PF | 19-NOV-2002; 2002US-0030611. |
| PX | |
| PR | 19-NOV-2002; 2002US-0030611. |
| PX | |
| PA | (ISIS-) ISIS PHARM INC. |
| PX | |
| PI | Chiang M, Doble KW; |
| PX | |
| DR | WPI; 2004-389192/36. |
| PX | |
| PT | New compounds, particularly oligonucleotides targeted to a nucleic acid encoding nldogen, useful for treating diseases associated with nldogen, |
| CC | e.g. Chediak-Higashi syndrome. |
| PX | |
| PS | Claim 1; SEQ ID NO 4; 91pp; English. |
| PX | |
| CC | The invention relates to antisense oligonucleotides which are targeted to, and inhibit the expression of, a nucleic acid molecule encoding nldogen. The antisense oligonucleotides are useful for treating a disease or condition associated with nldogen, such as Chediak-Higashi syndrome. |
| CC | They are also useful in research and diagnostics for modulating the expression of nldogen. The present sequence represents the gene encoding human nldogen. |
| CC | |
| SQ | Sequence 91352 BP; 23543 A; 20675 C; 20522 G; 26612 T; 0 U; 0 Other; |
| Query Match | 8.4%; Score 419; DB 12; Length 91352; |
| Best Local Similarity | 79.4%; Pred. No. 3e-66; |
| Matches 552; Conservative | 0; Mismatches 125; Indels 18; Gaps 4; |
| OY | 350 TTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGGAGTCGTCTGTGCGCCAGGCTGAG 409 |
| DB | 11334 TTATTATTAATTATTAATTATTAATTATTAATTATTAATTATTAATTATTAATTATTA 11339 |
| OY | 410 TGCAGTGCAGGAGTCTCGCTCACTCAAGCTCCGCTCCGGGTTCAAGCATTCTCT 469 |
| DB | 11394 TGCAGTGTGTGTGCTTGCTTAGCTCAAGCTCACCTCCGGGTTCCGGCATCTCT 11453 |
| OY | 470 GCCTCAAGCTCCCAAGTAGCTGGAGTCAAGCGGCCGCCACTAGCCCGGCTAATTTT 529 |
| DB | 11454 GCCTCAAGCTCCCAAGTAGCTGGAGTCAAGCACCCGCCACACGCCCGGCTAA-TTTT 11512 |
| OY | 530 TTGATATTTTAAAGTAGAGCGGGTTTACCGTTTAAAGCGGGATGGCTCGATCTCTGA 589 |
| DB | 11513 TTGATATTTTAAAGTAGAGCGGGTTTACCGTTTAAAGCGGGATGGCTCGATCTCTGA 11572 |
| OY | 590 CCTCGTATCCGCGCCGCTCGGCTCCCAAAGTCTGGAGTTACAAGCGCTGAGCCACCG 649 |
| DB | 11573 CCTCGTATCCACCACCTCGGCTCCCAAAGTCTGGAGTTACAAGCGCTGAGCCACCG 11632 |
| OY | 650 GCCCGGCGATGATCATCTTCTTGACTATGC---TGATGTGACAAAGTACTTAAGCATC 705 |
| DB | 11633 ACCGGGCGCTGAATTAATTTTAAAGCAATCCCAATATTTTATGCTTTGTAAATGTTTC 11692 |
| OY | 706 AGACTCAACCTTTAAATATG-CAGTTTGGCGCAGGACCCGTCATGCTCTGTAATTC 764 |
| DB | 11693 AGATGTGCTCAGAGAAGAAAAGATCTTTTGGCGGAGACAGATGGCTCATACCTGTAATTC 11752 |
| OY | 765 CAGCACTTTGGAGGAGAGAGTGGGTGAATCACTTGAGGCGAGGATTTGAACAGAGCT 824 |
| DB | 11753 CAGCACTTTGGAGGAGAGTGGGTGAATCACTTGAGGCGGATTTGAAGCTTCAGAGTTCCAGATGCTT 11812 |
| OY | 825 GGCCAACATGTGAAACTCTGTCTTTACTMAAAAAAAAAAAAAAAAAAAAAATACGCC 884 |
| DB | 11813 GGCCAACATGTGAAAACCTGTGTCACT-----AAAAATATCAAAATTAAGCC 11860 |
| OY | 885 GGGTGTGTGTGGGCACTCTGTAATCCAGCTATGCTGTGAAGCTGAGGCGAGAGTCAAC 944 |
| DB | 11861 AGATGTGTGTGGCACTCTGTAATCCAGCTATGCTGTGAAGCTGAGGCGAGAGTCAATTC 11920 |

| | | | |
|-----------|----------|--|--------|
| Oy | 340 | TCACTCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGCGGAGCTGCCTGCGCC | 399 |
| Db | 31712 | TCTATTATTAATTAATTTTTTTTTTTTTTTTTTTTGAGACGAGTCGTCTGTACCC | 317711 |
| Oy | 400 | CAGGCTGAGTGCAGTGGCGGAGATCTCGGCTCACTGAACCTCCGCTCCGGGTTCAAG | 459 |
| Db | 31772 | CAGGCTGAGTGCAGTGGCGGAGATCTTGCTCATCTGAACCTCACCTCCGGGTTCAAG | 318311 |
| Oy | 460 | CCATTCTTCCTGCTCAGCCTCCCAAGTAGTGGGACTACAGGCGGCCGCCCATAGCCCGG | 519 |
| Db | 31832 | CCATTCTTCCTGCTCAGCCTCCCAAGTAGTGGGATTAACAAGGCGGCCGCCCATAGCCCGG | 318931 |
| Oy | 520 | GCTAATTTTTTTTTTGAATTTTTTAGTAGAGACGGGGTTTTCAACGGTTTTAGCCGGATGACC | 579 |
| Db | 31892 | GCTAATTTTTTTTTTGAATTTTTTAGTAGAGACGGGGTTTTCAACGGTTTTAGCCGGATGACTC | 319531 |
| Oy | 580 | GATCTCCTGAACCTCGNATCCGCCCGCTCGGACTCCCAAAGTGTGGGATTAACAGCGCT | 639 |
| Db | 31952 | GATATCCTGAACCTTGATCCGCTCGCTCGGACTCCCAAAGTGTGGGATTAACAGCGCT | 320111 |
| Oy | 640 | GAGCCACCGCGCCCGGCATGATCATTTCTTGACTATGTGTGAACAATGACTTAA | 699 |
| Db | 32012 | GAGCCACCGGATCCGCGCGCTGTCTTGTATTAATTTTTAT-TCAAATAATTAATTGAG | 320701 |
| Oy | 700 | GCCATGACACTCACTACCTTTAAATATGCAAGTTGGGCGAGGCGACGCGTGCATGCTGT | 759 |
| Db | 32071 | GCAATCA-----GCTAGGCTTGGGCTCATGCTGTGC | 321021 |
| Oy | 760 | AATTCAGCACTTTGGAGGCAAGAGTGGTGAATCATCTTGAGGCCAGAGATTGAGAACC | 819 |
| Db | 32103 | AATCCAGCAC- TTGGGAGGCCAAGGCAAGGTGATCAGTTGAATGTGGGAGTTTGACATC | 321611 |
| Oy | 820 | AGCTGTGCGCAACATGTGGAACTCTGCTTTACTATAAAAAAAAAAAAAA-----A | 874 |
| Db | 32162 | AGCTGTGACAAACGTGGTAAACCCCACTCTTCAAAAAAAAAAAAAAAAAAACCA | 322211 |
| Oy | 875 | AAATTCAGCCGGGNGTGTGGGGGCAACCTGTATCCCAAGCTATGCTGAGGCTGAGGCA | 934 |
| Db | 32222 | AAAATTTGACCAAGCATTTTGGCGGGTCCCTGTAAATTCAGCTACCTGGGAGGCTGAGGCA | 322811 |
| Oy | 935 | CGABAGTCACTTGAACCTTGAGGCGGAGGTTGCAGTGGGCCGAGATCATCACCGCCC | 994 |
| Db | 32282 | GGAATATTGCTTGAACCTGGGAGGCGGAGGTTGCAAGTAGAGCTTAGATTGACCATGTGAC | 323411 |
| Oy | 995 | TCCAGCCTTGGGCGCAGAGCAAGCACTGTGTCTAAATTAATTAATAAACAAGAACAG | 1054 |
| Db | 32342 | TCCAGCCTTGGGCGCAGAGCAAGCAAGCCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAG | 324011 |
| Oy | 1055 | CAGTTTGTGTGACTTATGTTATATCTAAAA | 1084 |
| Db | 32402 | AAGGCGGTGTTGGTGGCTCAGCTGTGTA | 32431 |
| <hr/> | | | |
| RESULT 12 | | | |
| ADFe9167 | | | |
| ID | ADFe9167 | standard; cDNA; 216215 BP. | |
| XX | AC | ADFe9167; | |
| XX | DT | 12-FEB-2004 (first entry) | |
| XX | DE | Human MP53 nucleotide sequence SRQ ID NO:25. | |
| XX | KM | p53 pathway modulating agent; MP53; p53 modulator; cytosstatic; | |
| XX | KW | gene therapy; cancer; human; gene; ss. | |
| XX | OS | Homo sapiens. | |
| XX | PN | WO2003083047-A2. | |
| XX | PD | 09-OCT-2003. | |

| | |
|---------------------------|--|
| PF | 28-FEB-2003; 2003WO-US006025. |
| PR | 01-MAR-2002; 2002US-0361196P. |
| XX | (EXEL-) EXELIXIS INC. |
| PA | Belvin M, Francis-Lang H, Friedman L, Plowman GD, Heuer TS, Li D; |
| PI | Funkle RP; |
| XX | |
| XX | WPI: 2003-812540/76. |
| DR | P-PSDB; ADF69111. |
| XX | |
| PT | Identifying a candidate p53 pathway modulating agent for treating e.g., |
| PT | cancer by contacting an assay system comprising a MP53 polypeptide or |
| PT | nucleic acid with a test agent and detecting a test agent-biased |
| XX | activity. |
| XX | |
| PS | Example; SEQ ID NO 25; 406pp; English. |
| XX | |
| CC | The present invention describes a method for identifying a candidate p53 |
| CC | pathway modulating agent, which comprises: (a) providing an assay system |
| CC | comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its |
| CC | fragment or derivative; (b) contacting the assay system with a test agent |
| CC | under conditions where the system provides a reference activity except in |
| CC | the presence of the test agent; and (c) detecting a test agent-biased |
| CC | activity, where a difference between the test agent-biased activity and |
| CC | the reference activity identifies the test agent as a candidate p53 |
| CC | pathway modulating agent. Also described: (1) modulating the p53 pathway |
| CC | of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3) |
| CC | diagnosing a disease in a patient. MP53 has cytosolic activity, and can |
| CC | be used in gene therapy. The method is useful for identifying a candidate |
| CC | p53 pathway modulating agent for preparing a composition for diagnosing |
| CC | or treating e.g., cancer. The present sequence encodes a human MP53 |
| CC | protein, which is used in the exemplification of the present invention. |
| XX | |
| SEQ | Sequence 216215 BP; 64591 A; 41616 C; 43966 G; 66042 T; 0 U; 0 Other; |
| Query Match | 8.3%; Score 417.4; DB 10; Length 216215; |
| Best Local Similarity | 77.8%; Pred. No. 6, 4e-66; |
| Matches 544; Conservative | 0; Mismatches 146; Indels 9; Gaps 3; |
| QY | 347 CTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCGTCTGTGCCAGGCTG 406 |
| DB | 117634 CTTTCTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGCGCTTGCTGTGCACCAGGCTG 117693 |
| QY | 407 GAGTGCATGGGGGAGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCATTCT 466 |
| DB | 117694 GAGTGCATGGGGGAGATCTCAACTCACTGCAACTCTTGAAGTTCAAGCTATTCT 117753 |
| QY | 467 CCTGCTCAGCTCTCCAGTAGCTGGAGCTACAGCGCGCCGACATAAGCCCGGCTAATT 526 |
| DB | 117754 TCTGCTCAGCTCTCCAGTAGCTGGAGCTACAGCGCTGACCCCATGCCAGCTTA-- 117811 |
| QY | 527 TTTTGTATTTTAAAGAGACGGGGTTTACCGTTTAAAGCCGGAGTGGCTCTGATCTCC 586 |
| DB | 117812 TTTTGTATTTTAAAGAGAGGGGTTTACCATATGCGCAGGGTGTCTTGAATCC 117871 |
| QY | 587 TGACCTCGTGAATCCGCGCGCTCGGCTCCCAAAGTGTGGGATTAAGAAGCGAGCCAC 646 |
| DB | 117872 TGACCTGTGAATCTCGCTCGCTCAAGCTCCCAAAGTGTGGGATTAAGAAGCGAGCCAC 117931 |
| QY | 647 CGCGCCGCGCATGATCATCTTCTTGACTATCTGATGAGACAAGTAAAGCATCA 706 |
| DB | 117932 TGTGACTGGCCCTCATCCACTTCTTACACACCAACCTTTTAG-----GTGTATT 117985 |
| QY | 707 GACTCTACCTTTAATATATGAGTTTGGCCAGGACCGGTGCTCATGCTGTAATTCCA 766 |
| DB | 117986 TTTCCTCATGTAAAGAAAATCTTTCCGCGCAGGACGGGTGCTCACGCTGTAATCCA 118045 |
| QY | 767 GCACCTTTGGAGAGAGAGTGGGTGAATCATCTTAGAGCCAGAGATTGAGACAGGCTGG 826 |
| DB | 118046 GCACCTTTGGAGAGAGAGTGGGTGAATCATCTTAGAGCCAGAGATTGAGACAGGCTGG 118105 |


```
QY 649 CGCCCGGCATGATCATCTTCTTGAATGATGACAGTAAAGCCATCAGA 708
    |||||
Db 15449 CACCCGCC-----GATGTTGCACTCTTAAAGTGAACAAATCCAGCTTTA-- 15501
    |||||
QY 709 CTCACCCCTTAAATATGAGTTTGAGCCAGGACCGTGGCTCATGCTGTAAATCCAGC 768
    |||||
Db 15502 -----TAAATATATATAGCTGGGCTGGGCGCGGTGGCTCACACTGTAAATCCAGC 15552
    |||||
QY 769 ACTTTGGAGGCGAGAGGTGGGTGAATCATCTTGAGGCCAGAGTTTGAAGCAGCCTGGCC 828
    |||||
Db 15553 AGTTTGGAGGCGCAGGCGAGGTGAATCATCTTGAGCTGGGTGTTCAAGACCAAGCTTGGCC 15612
    |||||
QY 829 AACATGGTAAACTGTCTTCTTAACTAATAAAAAAAAAAAAAAAAAATCAGCCGGGT 888
    |||||
Db 15613 AACATGGTAAACGCTGTCTTACT-----AAAAATACAAAAATTAGCCAGGT 15660
    |||||
QY 889 GTGCTGGGGGCAACCTGTAAATCCAGCTATGCTGAGAGGCTGAGGCAAGAGTCACTTGA 948
    |||||
Db 15661 GTGCTGGGCAACACTGTAAATCCAGCTACTTGGAGGCTGAAGCAAGAAATCACTTGA 15720
    |||||
QY 949 ACCCTGAGGCGGAGGTGAGTGGGCGGATCAGATCACGCCCTCCAGCCTGGGCGA 1008
    |||||
Db 15721 ACCCGGAGGCGAGAGGTGAGTGGCGGAGATTATGCCACTGCACCTCCACTGGGAGTGA 15780
    |||||
QY 1009 CAGAGCAAGACTGTCTCAATTAATAATAAACAACGACAGCAGT 1058
    |||||
Db 15781 AAGAGCAAGACTGTCTCAAAACATTAAAAAAATAAATAATAAAGT 15830
    |||||
```

Search completed: February 25, 2005, 20:14:49
Job time : 2552.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 18:05:22 ; Search time 8524.79 Seconds
(without alignments)
11373.754 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gcaacacagagagcccccag.....tcgcagggggaacatttca 2001

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of Hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|---------------------|
| 1 | 2001 | 100.0 | 11570 | 6 | CQ788219 |
| 2 | 2001 | 100.0 | 11570 | 6 | CQ874885 |
| 3 | 2001 | 100.0 | 11570 | 6 | CQ874885 Sequence |
| 4 | 2001 | 100.0 | 11570 | 6 | CQ874885 Sequence |
| 5 | 2001 | 100.0 | 107487 | 9 | AF135028 Homo sapi |
| 6 | 2001 | 100.0 | 230000 | 2 | AF243527 Homo sapi |
| 7 | 1968.6 | 98.4 | 217346 | 2 | AC027602 Homo sapi |
| 8 | 1977.2 | 96.3 | 200792 | 2 | AC130782 Pan trogl |
| 9 | 286.4 | 14.3 | 142003 | 9 | AC098800 Homo sapi |
| 10 | 276.8 | 13.8 | 177877 | 9 | AC133961 Homo sapi |
| 11 | 274.4 | 13.2 | 138411 | 9 | HSBJ11H19 Human DNA |
| 12 | 264 | 13.2 | 161014 | 9 | AC087591 Homo sapi |
| 13 | 264 | 13.2 | 177876 | 2 | AC025163 Homo sapi |
| 14 | 262.4 | 13.1 | 92882 | 9 | AL355815 Human DNA |
| 15 | 261.8 | 13.1 | 85304 | 9 | HS227L5 Human DNA |
| 16 | 258.4 | 12.9 | 177562 | 9 | HS269M15 Human DNA |
| 17 | 257.8 | 12.9 | 70356 | 9 | AL353794 Human DNA |
| 18 | 257.8 | 12.9 | 165330 | 2 | AC016011 Homo sapi |
| 19 | 255.8 | 12.8 | 154664 | 9 | AC025362 Homo sapi |

| | | | | | |
|----|-------|------|--------|---|---------------------|
| 20 | 255.8 | 12.8 | 160426 | 2 | AC087697 Homo sapi |
| 21 | 255.8 | 12.8 | 162354 | 2 | AC073299 Homo sapi |
| 22 | 255.8 | 12.8 | 179219 | 9 | AC016489 Homo sapi |
| 23 | 255.2 | 12.8 | 74705 | 9 | AC116339 Homo sapi |
| 24 | 255.2 | 12.8 | 161133 | 2 | AC019295 Homo sapi |
| 25 | 255.2 | 12.8 | 163690 | 2 | AC083781 Homo sapi |
| 26 | 255.2 | 12.8 | 195476 | 9 | AC010259 Homo sapi |
| 27 | 253.8 | 12.7 | 136649 | 9 | HSBJ635E8 Human DNA |
| 28 | 253.4 | 12.7 | 160420 | 9 | AC022023 Homo sapi |
| 29 | 253.4 | 12.7 | 162609 | 9 | AC021171 Homo sapi |
| 30 | 253.4 | 12.7 | 179009 | 9 | AL355543 Human DNA |
| 31 | 253.2 | 12.7 | 173840 | 9 | AC092445 Homo sapi |
| 32 | 253.2 | 12.7 | 206819 | 2 | AC016029 Homo sapi |
| 33 | 253 | 12.6 | 193772 | 9 | AC007073 Homo sapi |
| 34 | 252.6 | 12.6 | 52216 | 6 | BD012166 Isolation |
| 35 | 252.6 | 12.6 | 52216 | 6 | BD128690 Method fo |
| 36 | 252.6 | 12.6 | 163453 | 9 | AC003982 Homo sapi |
| 37 | 252.6 | 12.6 | 163453 | 9 | AC003982 Homo sapi |
| 38 | 252.6 | 12.6 | 169544 | 9 | AC008573 Homo sapi |
| 39 | 252.6 | 12.6 | 179665 | 2 | AC010246 Homo sapi |
| 40 | 252.6 | 12.6 | 197810 | 9 | AC098071 Homo sapi |
| 41 | 252.6 | 12.6 | 220384 | 2 | AC078926 Homo sapi |
| 42 | 252.6 | 12.6 | 253038 | 2 | AC008930 Homo sapi |
| 43 | 251.4 | 12.6 | 33768 | 9 | AC078783 Homo sapi |
| 44 | 250.6 | 12.5 | 103585 | 9 | AC113614 Homo sapi |
| 45 | 250.6 | 12.5 | 180939 | 2 | AL596206 Homo sapi |

ALIGNMENTS

| | | | | | | |
|----------------------------|---|---|----------|-----|--------|-----------------|
| RESULT 1 | CQ788219 | Sequence 2 from Patent WO2004021008. | 11570 bp | DNA | linear | PAT 24-MAR-2004 |
| LOCUS | CQ788219 | Sequence 2 from Patent WO2004021008. | | | | |
| DEFINITION | CQ788219 | Sequence 2 from Patent WO2004021008. | | | | |
| ACCESSION | CQ788219 | Sequence 2 from Patent WO2004021008. | | | | |
| VERSION | CQ788219.1 | GI:45723068 | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |
| | Homo sapiens (human) | | | | | |
| | Homo sapiens | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | 1 | Diamandis E.P. | | | | |
| TITLE | | Methods for detecting breast and ovarian cancer | | | | |
| JOURNAL | | Patent: WO 2004021008-A 2 11-MAR-2004; | | | | |
| | | MOUNT SINAI HOSPITAL (CA); Yousef, George (CA) | | | | |
| FEATURES | | | | | | |
| source | | 1. 11570 | | | | |
| | | /organism="Homo sapiens" | | | | |
| | | /mol_type="unassigned DNA" | | | | |
| | | /db_xref="taxon:9606" | | | | |
| ORIGIN | | | | | | |
| Query Match | | 100.0%; Score 2001; DB 6; Length 11570; | | | | |
| Best Local Similarity | | 100.0%; Pred. No. 0; | | | | |
| Matches 2001; Conservative | | 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY | 1 | GACACACACAGAGCCCAAGAGTGTCCAGGTTCTTCTTATACGACCACTCTGTC | 60 | | | |
| DB | 6000 | GACACACACAGAGCCCAAGAGTGTCCAGGTTCTTCTTATACGACCACTCTGTC | 6059 | | | |
| QY | 61 | CGCCTTCATCTTCTTCCATCTTCAATGTGTTCCGTTTGAACAGTCCCTTAAG | 120 | | | |
| DB | 6060 | CGCCTTCATCTTCTTCCATCTTCAATGTGTTCCGTTTGAACAGTCCCTTAAG | 6119 | | | |
| QY | 121 | TCCCTCAGCTCTTGAATATCAGGCTGTAGTCAGAAAGGTGCGAGATCTTACCGCA | 180 | | | |
| DB | 6120 | TCCCTCAGCTCTTGAATATCAGGCTGTAGTCAGAAAGGTGCGAGATCTTACCGCA | 6179 | | | |
| QY | 181 | GACAGATAGATGACACCATGTTCTGCGCCGAGTACAAAGCAGGTAGAGACTCTGCCAGG | 240 | | | |
| DB | 6180 | GACAGATAGATGACACCATGTTCTGCGCCGAGTACAAAGCAGGTAGAGACTCTGCCAGG | 6239 | | | |

| | | | |
|----|------|---|------|
| QY | 241 | GGAGGACACCTCTCTTTTATTCAGAGATACACCTGAGTGCACCTGCTATCAATGGAC | 300 |
| Db | 6240 | TGAGGACACCTCTCTTTTATTCAGAGATACACCTGAGTGCACCTGCTATCAATGGAC | 6299 |
| QY | 301 | GTTCGCAATTCCTGGAATCCAGCAATTCGACAGACAGTCAAGAACCTCTGTTCTACAGA | 360 |
| Db | 6300 | GTTCGCAATTCCTGGAATCCAGCAATTCGACAGACAGTCAAGAACCTCTGTTCTACAGA | 6359 |
| QY | 361 | GCTCATACCCATGAGTATGAGTGTGTTTACTAGAAATATATGCTGAGCTGTTATGTCATTTCC | 420 |
| Db | 6360 | GCTCATACCCATGAGTATGAGTGTGTTTACTAGAAATATATGCTGAGCTGTTATGTCATTTCC | 6419 |
| QY | 421 | CAGTTTTTTAGTAGGACACATTTAAACAGSTAAAAAAGCGCTGGGCGACAGTGCACACCT | 480 |
| Db | 6420 | CAGTTTTTTAGTAGGACACATTTAAACAGSTAAAAAAGCGCTGGGCGACAGTGCACACCT | 6479 |
| QY | 481 | GTAATCCAGCACTTTGGAGGCTGAGGACAGGACAGATCACCTTTGTGACGAGTTTGAGA | 540 |
| Db | 6480 | GTAATCCAGCACTTTGGAGGCTGAGGACAGGACAGATCACCTTTGTGACGAGTTTGAGA | 6539 |
| QY | 541 | CTAGCTGCGGCAACATGGGCGAACTCTGCTCTATAAAAAAATCAAAATATATACCTGGCC | 600 |
| Db | 6540 | CTAGCTGCGGCAACATGGGCGAACTCTGCTCTATAAAAAAATCAAAATATATACCTGGCC | 6599 |
| QY | 601 | ATGTGGCGGGGCGCCTGTATCTCAGCTCAGGAGGCGCAGACACAAAGATCACTTAA | 660 |
| Db | 6600 | ATGTGGCGGGGCGCCTGTATCTCAGCTCAGGAGGCGCAGACACAAAGATCACTTAA | 6659 |
| QY | 661 | ACCCAGAGAGTGGAGGTGGCAGTGAATGTCGATCTCATCTCATCTCCAACTGGGAGAC | 720 |
| Db | 6660 | ACCCAGAGAGTGGAGGTGGCAGTGAATGTCGATCTCATCTCATCTCCAACTGGGAGAC | 6719 |
| QY | 721 | AGAGTGAACACTTGTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGSTGAAG | 780 |
| Db | 6720 | AGAGTGAACACTTGTGTCTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGSTGAAG | 6779 |
| QY | 781 | TTAATCTTAAATPAAACCAATGTATCCCAATACATCAATTCGAAAGTAAATATPAA | 840 |
| Db | 6780 | TTAATCTTAAATPAAACCAATGTATCCCAATACATCAATTCGAAAGTAAATATPAA | 6839 |
| QY | 841 | ACAATTAATGAATGAGATACCTTACATCTTTCTGTTTCAATPATAAGCTTTGAAGT | 900 |
| Db | 6840 | ACAATTAATGAATGAGATACCTTACATCTTTCTGTTTCAATPATAAGCTTTGAAGT | 6899 |
| QY | 901 | GAGTATATATGTATGTGAACAGCACATCTCAATTTGGACATGACTACATTTCAAGTGTCTC | 960 |
| Db | 6900 | GAGTATATATGTATGTGAACAGCACATCTCAATTTGGACATGACTACATTTCAAGTGTCTC | 6959 |
| QY | 961 | AGTACCCACATGTGGCTAGCAGTTATCTGTATTTGATGGCACAGGATCTTAGAGGAAAGATC | 1020 |
| Db | 6960 | AGTACCCACATGTGGCTAGCAGTTATCTGTATTTGATGGCACAGGATCTTAGAGGAAAGATC | 7019 |
| QY | 1021 | AGGGCTGTTTGTATGTGGGACAGGTTGTGCATGCACTAAAGATACCATATCTAATAGG | 1080 |
| Db | 7020 | AGGGCTGTTTGTATGTGGGACAGGTTGTGCATGCACTAAAGATACCATATCTAATAGG | 7079 |
| QY | 1081 | GGCACTCCGTTTACAGATGTCAAGTTTGGCAGTTTCAAGCGTGTGTAGTTAATGTCTC | 1140 |
| Db | 7080 | GGCACTCCGTTTACAGATGTCAAGTTTGGCAGTTTCAAGCGTGTGTAGTTAATGTCTC | 7139 |
| QY | 1141 | TTGTTTCAACAAAATCTGTAAATATGACAGTTTCTAGCAAGTGTGCTGTAAATATCTTTGA | 1200 |
| Db | 7140 | TTGTTTCAACAAAATCTGTAAATATGACAGTTTCTAGCAAGTGTGCTGTAAATATCTTTGA | 7199 |
| QY | 1201 | GGAAAGAAAGAGAAATCTGTGTAGTTATTTTTCACAGAAATATTTATACAGGGGATTA | 1260 |
| Db | 7200 | GGAAAGAAAGAGAAATCTGTGTAGTTATTTTTCACAGAAATATTTATACAGGGGATTA | 7259 |
| QY | 1261 | ATTGCAAGCTGCTGGAAAGGGCTGGAGAACAAAGTTAAAAATATAAAAATCTCTGTGCTC | 1320 |
| Db | 7260 | ATTGCAAGCTGCTGGAAAGGGCTGGAGAACAAAGTTAAAAATATAAAAATCTCTGTGCTC | 7319 |

| | | | |
|----|------|--|------|
| QY | 1321 | TAAGATCTGCATAAATATGGGCAATTTGACAGAGTGTAAAGCTTAAACCCCAAAATPAAAC | 1380 |
| Db | 7320 | AAGAAATCTGCATAAATATGGGCAATTTGACAGAGTGTAAAGCTTAAACCCCAAAATPAAAC | 7379 |
| QY | 1361 | ATGTTTTAGATAGTAATAACATAAGGCCAATATTCAAAAAGTGTCAAGGGAGCCTC | 1440 |
| Db | 7380 | ATGTTTTAGATAGTAATAACATAAGGCCAATATTCAAAAAGTGTCAAGGGAGCCTC | 7439 |
| QY | 1441 | CTTGAGAGGTGGCAATTTGACACAGAAATGAAATGACCAAGAAGCTTAACTCGTGAAGT | 1500 |
| Db | 7440 | CTTGAGAGGTGGCAATTTGACACAGAAATGAAATGACCAAGAAGCTTAACTCGTGAAGT | 7499 |
| QY | 1501 | TTAAGGGGAAAAGAAAAGCAGTCGCAAGGCCCTGAGGACAGTAAGAAATTTGGCTGATTC | 1560 |
| Db | 7500 | TTAAGGGGAAAAGAAAAGCAGTCGCAAGGCCCTGAGGACAGTAAGAAATTTGGCTGATTC | 7559 |
| QY | 1561 | AAAGAGAGAGAGAAACCAATGCACTGGAGAACAAAGTGGGGGCAACAGTAGAAAGTG | 1620 |
| Db | 7560 | AAAGAGAGAGAGAAACCAATGCACTGGAGAACAAAGTGGGGGCAACAGTAGAAAGTG | 7619 |
| QY | 1621 | ACGCTGAGAGTGTAGGCGAGGGGCGCAATGCTGTGCAAGATTTCTTGGTCACCAACAGAA | 1680 |
| Db | 7620 | ACGCTGAGAGTGTAGGCGAGGGGCGCAATGCTGTGCAAGATTTCTTGGTCACCAACAGAA | 7679 |
| QY | 1681 | GCTTCCTATGTTCTAATGAGAACTGTATCTGTTGAGGAGACAGAAATTTAAATCAAAC | 1740 |
| Db | 7680 | GCTTCCTATGTTCTAATGAGAACTGTATCTGTTGAGGAGACAGAAATTTAAATCAAAC | 7739 |
| QY | 1741 | TGTTACATCAACACAGACACCTTCTCTGTATTCAGAGTCCCAAGGAATCTGAAAGACGTA | 1800 |
| Db | 7740 | TGTTACATCAACACAGACACCTTCTCTGTATTCAGAGTCCCAAGGAATCTGAAAGACGTA | 7799 |
| QY | 1801 | AGTTAACAAAGCTCTCATTTAGCAGGAGTGTGTTTCAACAGTAGTTAGGAAGCTGGGAGTT | 1860 |
| Db | 7800 | AGTTAACAAAGCTCTCATTTAGCAGGAGTGTGTTTCAACAGTAGTTAGGAAGCTGGGAGTT | 7859 |
| QY | 1861 | CAGAGATCTCCACATCCCATGGCTATGAAAAGCTCCCCCAAAATTTGTAACAACTGACAA | 1920 |
| Db | 7860 | CAGAGATCTCCACATCCCATGGCTATGAAAAGCTCCCCCAAAATTTGTAACAACTGACAA | 7919 |
| QY | 1921 | ATGCACAACCTCCACAGCTCTCCCATTTCTTCTGTGCTCCCTGGGTGTGGGGGGGTGGG | 1980 |
| Db | 7920 | ATGCACAACCTCCACAGCTCTCCCATTTCTTCTGTGCTCCCTGGGTGTGGGGGGGTGGG | 7979 |
| QY | 1981 | TTGCGAGGGGAAAACCTTTTA 2001 | |
| Db | 7980 | TTGCGAGGGGAAAACCTTTTA 8000 | |

| | | | | |
|-------------|---|-------------|-------|-----------------|
| RESULT 2 | | | | |
| CO874885 | CO874885 | 11570 bp | DNA | linear |
| LOCUS | CO874885 | 11570 bp | DNA | PAT 27-SEP-2004 |
| DEFINITION | Sequence 3 from Patent WO2004077060. | | | |
| ACCESSION | CO874885 | | | |
| VERSION | CO874885.1 | GI:52748035 | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | Homo sapiens (human) | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | | |
| REFERENCE | 1 | | | |
| AUTHORS | Diamandis, E. P. and Petraki, C. | | | |
| TITLE | Assay for detection of renal cell carcinoma | | | |
| JOURNAL | Patent: WO 2004077060-A 3 10-SEP-2004; | | | |
| | Mount Sinai Hospital (CA) | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..11570 | | | |
| | /organism="Homo sapiens" | | | |
| | /mol_type="unassigned DNA" | | | |
| | /db_xref="taxon:9606" | | | |
| ORIGIN | | | | |
| Query Match | 100.0%; | Score 2001; | DB 6; | Length 11570; |

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 GACAAACCAAGAGCCCCCAAGGTGATCCAGGTTCTTTGTAATCCAGCCATCTTGC 60
Db 6000 GACAAACCAAGAGCCCCCAAGGTGATCCAGGTTCTTTGTAATCCAGCCATCTTGC 6059
Qy 61 CGGCTTCATCTTCTCAGCTCTCATTTGTTCTGTTGACGTGCACTTCCCTAAG 120
Db 6060 CGGCTTCATCTTCTCAGCTCTCATTTGTTCTGTTGACGTGCACTTCCCTAAG 6119
Qy 121 TCTCCAGGCTTGAATATGAGCTGTGAATGAGAAAGGTGAGATGCTTACCGA 180
Db 6120 TCTCCAGGCTTGAATATGAGCTGTGAATGAGAAAGGTGAGATGCTTACCGA 6179
Qy 181 GACAGATGATGACACCATGTTCTGCGCCGCTGACAAAGCAGTAAAGTCTGCGAG 240
Db 6180 GACAGATGATGACACCATGTTCTGCGCCGCTGACAAAGCAGTAAAGTCTGCGAG 6239
Qy 241 TGAAGACACCTCTTATTCAGAGATPACACTGAGTGCAACTGAGTAACTGAGC 300
Db 6240 TGAAGACACCTCTTATTCAGAGATPACACTGAGTGCAACTGAGTAACTGAGC 6299
Qy 301 GTTGCCAAATTTGAGATCCAGCAATTGCCAAAGACAGTACAGACCCCTGTTCAAGA 360
Db 6300 GTTGCCAAATTTGAGATCCAGCAATTGCCAAAGACAGTACAGACCCCTGTTCAAGA 6359
Qy 361 GCTCATACCTAGAGATGAGTGTGTTAGTAAATTAATGCTAGCTGTTATGCTATTC 420
Db 6360 GCTCATACCTAGAGATGAGTGTGTTAGTAAATTAATGCTAGCTGTTATGCTATTC 6419
Qy 421 CAGTTTTTATGAGGCACTTAAACAGGTAAACAGGTGAGGAGGAGTGCACACCT 480
Db 6420 CAGTTTTTATGAGGCACTTAAACAGGTAAACAGGTGAGGAGGAGTGCACACCT 6479
Qy 481 GTATCCACAGCACTTTGGAGGCTGAGGAGGAGAGTCACTTTGATCAGAGTTTGA 540
Db 6480 GTATCCACAGCACTTTGGAGGCTGAGGAGGAGAGTCACTTTGATCAGAGTTTGA 6539
Qy 541 CTAGCTGCGCAACATGCGCAAACTGTGTCTTAAATAAATAAATTAAGCTGCG 600
Db 6540 CTAGCTGCGCAACATGCGCAAACTGTGTCTTAAATAAATAAATTAAGCTGCG 6599
Qy 601 ATGTGCGGCGCGCTGTAATCTCAGCTGCTCAGGAGGCGGAGCACAAGATCACTTA 660
Db 6600 ATGTGCGGCGCGCTGTAATCTCAGCTGCTCAGGAGGCGGAGCACAAGATCACTTA 6659
Qy 661 ACCCAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCAACCTGAGAG 720
Db 6660 ACCCAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCAACCTGAGAG 6719
Qy 721 AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAGAAAAAGAGTGAAG 780
Db 6720 AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAGAAAAAGAGTGAAG 6779
Qy 781 TTAACTTAAATTAACCAATGTAATCCAAATTAATCAATTTCAAGTAAATTAATAA 840
Db 6780 TTAACTTAAATTAACCAATGTAATCCAAATTAATCAATTTCAAGTAAATTAATAA 6839
Qy 841 ACAATTAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Db 6840 ACAATTAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6899
Qy 901 GAGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 6900 GAGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6959
Qy 961 AGTAGCCACATGTTGGCTAGCACTTAATGATGATGATGATGATGATGATGATGAT 1020
Db 6960 AGTAGCCACATGTTGGCTAGCACTTAATGATGATGATGATGATGATGATGATGAT 7019
Qy 1021 AGGAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
```

```
Db 7020 AGGAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7079
Qy 1081 GGCATCCGCTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 7080 GGCATCCGCTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7139
Qy 1141 TTGTTTCAACAAATCTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 7140 TTGTTTCAACAAATCTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 7199
Qy 1201 GGAAGAAAAAGAAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 7200 GGAAGAAAAAGAAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 7259
Qy 1261 ATTGCAAGCTGCTGGAAGGCTGAGAGGAAACAAGTAAATTAATTAATTAATTAAT 1320
Db 7260 ATTGCAAGCTGCTGGAAGGCTGAGAGGAAACAAGTAAATTAATTAATTAATTAAT 7319
Qy 1321 AAGAACTGCAATTAATAGGCAATTTCAAGAGTGTGTAAGTAAACCCCAATTAATAC 1380
Db 7320 AAGAACTGCAATTAATAGGCAATTTCAAGAGTGTGTAAGTAAACCCCAATTAATAC 7379
Qy 1381 ATGGTTTATGATGATGATTAACAATTAAGGCAATTAATTAATTAATTAATTAAT 1440
Db 7380 ATGGTTTATGATGATGATTAACAATTAAGGCAATTAATTAATTAATTAATTAAT 7439
Qy 1441 CTTGAGAGGTGCTATTTGAGCAGAGATGAGATGAGACAAAGAGCTTAATCTGTAAGT 1500
Db 7440 CTTGAGAGGTGCTATTTGAGCAGAGATGAGATGAGACAAAGAGCTTAATCTGTAAGT 7499
Qy 1501 TTAAGGGGAAAAAAGGACGTCGAAAGGCTGAGAGGAGTAAAGTATGCTGATTC 1560
Db 7500 TTAAGGGGAAAAAAGGACGTCGAAAGGCTGAGAGGAGTAAAGTATGCTGATTC 7559
Qy 1561 AAAGAAAGAGAGGAAACCAATGCACTGAGAAACAAAGTGGGGCAACAGTAAAGT 1620
Db 7560 AAAGAAAGAGAGGAAACCAATGCACTGAGAAACAAAGTGGGGCAACAGTAAAGT 7619
Qy 1621 ACCGTGAGGTGATGAGCAGGAGGAGATGCTCTGCAAGTATTTCTTGCTCAACACAG 1680
Db 7620 ACCGTGAGGTGATGAGCAGGAGGAGATGCTCTGCAAGTATTTCTTGCTCAACACAG 7679
Qy 1681 GCTTCCCTATGTTCTTAATGGAAGCTGATCTGTTGAGGAAAGACAGATTAATCAAC 1740
Db 7680 GCTTCCCTATGTTCTTAATGGAAGCTGATCTGTTGAGGAAAGACAGATTAATCAAC 7739
Qy 1741 TGTTAATCAACAGCAACCTTCTGTAATTCAGGCTCCCAAGGATCTAAGAGAGT 1800
Db 7740 TGTTAATCAACAGCAACCTTCTGTAATTCAGGCTCCCAAGGATCTAAGAGAGT 7799
Qy 1801 AGTTAACAGCTCTCATTTAGCAGGAGTGTGTTCAACAGTAAAGTAAAGCTGGGAT 1860
Db 7800 AGTTAACAGCTCTCATTTAGCAGGAGTGTGTTCAACAGTAAAGTAAAGCTGGGAT 7859
Qy 1861 CAGAGTATCCAGTCCCATGCTATGATGAAAGCTCCCAATTTGTAACAACTGACAA 1920
Db 7860 CAGAGTATCCAGTCCCATGCTATGATGAAAGCTCCCAATTTGTAACAACTGACAA 7919
Qy 1921 ATGCAACACCTCCAGCTCTCCCAATTTCTTCTGTTGCTGAGGTTGGGGGGTGG 1980
Db 7920 ATGCAACACCTCCAGCTCTCCCAATTTCTTCTGTTGCTGAGGTTGGGGGGTGG 7979
Qy 1981 TTGCGAGGGGAAAACTTTTA 2001
Db 7980 TTGCGAGGGGAAAACTTTTA 8000
```

RESULT 3
C0874961 11570 bp DNA linear PART 27-SEP-2004
LOCUS C0874961
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION C0874961
VERSION C0874961.1 GI:52748060

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Diamondis, E. P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
source 1. .11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2001; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCAAGAGGCCCAAGAGAGTCCAGGTTCTTCTTGAACCCATCTCTGC 60
DB 6000 GACACCAAGAGGCCCAAGAGAGTCCAGGTTCTTCTTGAACCCATCTCTGC 6059
QY 61 CGCCTTCACATCTTCTCCTCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 120
DB 6060 CGCCTTCACATCTTCTCCTCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 6119
QY 121 TCTCCAGTCTGGAATATCAGCGTGTAACTGACAAAAGTCCAGAGATCTTACCCGA 180
DB 6120 TCTCCAGTCTGGAATATCAGCGTGTAACTGACAAAAGTCCAGAGATCTTACCCGA 6179
QY 181 GACAGATAGATGACACATGTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGCGAG 240
DB 6180 GACAGATAGATGACACATGTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGCGAG 6239
QY 241 TGAGGACACCTCTCTTATTCAGCAGATACACACTGAGTCCAACTCGTTAATGAGAG 300
DB 6240 TGAGGACACCTCTCTTATTCAGCAGATACACACTGAGTCCAACTCGTTAATGAGAG 6299
QY 301 GTTGCCAAATTCGAGAAATCAGCAATTCGCAAGACATGTCAGAGCCCTGTTCTCAAGA 360
DB 6300 GTTGCCAAATTCGAGAAATCAGCAATTCGCAAGACATGTCAGAGCCCTGTTCTCAAGA 6359
QY 361 GCTCATACCTTAGAGTATGTTAGTGAATAATGCTGAGTGTATGCTCAATTC 420
DB 6360 GCTCATACCTTAGAGTATGTTAGTGAATAATGCTGAGTGTATGCTCAATTC 6419
QY 421 CAGTTTTTTAGTACCAATTAACAGGTAAAAAGGCTGGCGAGTGGCTCAACCT 480
DB 6420 CAGTTTTTTAGTACCAATTAACAGGTAAAAAGGCTGGCGAGTGGCTCAACCT 6479
QY 481 GTTAATCCAGCACTTTGGAGGCTGAGGCGAGCAATCACCTTTGGTCAAGAGTTGAGA 540
DB 6480 GTTAATCCAGCACTTTGGAGGCTGAGGCGAGCAATCACCTTTGGTCAAGAGTTGAGA 6539
QY 541 CTAGCCCTGGCCAACTGGGCAAACTGTTCTTAAAAAAATTAAGGCTGGG 600
DB 6540 CTAGCCCTGGCCAACTGGGCAAACTGTTCTTAAAAAAATTAAGGCTGGG 6599
QY 601 ATGGTGGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGGAGACAAAGATCACTTAA 660
DB 6600 ATGGTGGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGGAGACAAAGATCACTTAA 6659
QY 661 ACCCAAGAGGTGAGGTGAGTGAAGTGAATGTCCTCACTCACTCAAGCTGGAGAC 720
DB 6660 ACCCAAGAGGTGAGGTGAGTGAAGTGAATGTCCTCACTCAAGCTGGAGAC 6719
QY 721 AGAGTGAACCTTTGTCTCAAAAAGAAAAAAACAAGTAAAAAGAAACAGGTGAAG 780
DB 6720 AGAGTGAACCTTTGTCTCAAAAAGAAAAAAACAAGTAAAAAGAAACAGGTGAAG 6779

QY 781 TTAACCTTAATAACCAATGATATCCCAATTAACATCATTTCAAGTATTAATAA 840
DB 6780 TTAACCTTAATAACCAATGATATCCCAATTAACATCATTTCAAGTATTAATAA 6839
QY 841 ACAATTAATAAGATTAATCTTATCATTTCTTTCTGTTTCAATTAAGCTTTGAAAGT 900
DB 6840 ACAATTAATAAGATTAATCTTATCATTTCTTTCTGTTTCAATTAAGCTTTGAAAGT 6899
QY 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTGATCAATTTGAGGCTC 960
DB 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTGATCAATTTGAGGCTC 6959
QY 961 AGTAGCCATGTGGTGTACAGTATCTGTATTTGATGCGACCGATCTAGAGGAAAGATC 1020
DB 6960 AGTAGCCATGTGGTGTACAGTATCTGTATTTGATGCGACCGATCTAGAGGAAAGATC 7019
QY 1021 AGGGCTGTTTGTATGCTTGGGAGGTTGTGACCTGATTAAGATATCTAATAGG 1080
DB 7020 AGGGCTGTTTGTATGCTTGGGAGGTTGTGACCTGATTAAGATATCTAATAGG 7079
QY 1081 GGCACCTCGGTGTAAGATGTCAGTGTGTTGGCAGTTTTCAGGGGTGTAGTTAAGTTC 1140
DB 7080 GGCACCTCGGTGTAAGATGTCAGTGTGTTGGCAGTTTTCAGGGGTGTAGTTAAGTTC 7139
QY 1141 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTCTGTATTAATCTTGA 1200
DB 7140 TTGTTTCAACAAATCTGTATATGACAGTTTCTTAGCAAGTCTGTATTAATCTTGA 7199
QY 1201 GGAAGAAAAAGAAATCTGTAGGTAATTTTTCAGAGAAATTTTAAATACAGGGGATTA 1260
DB 7200 GGAAGAAAAAGAAATCTGTAGGTAATTTTTCAGAGAAATTTTAAATACAGGGGATTA 7259
QY 1261 ATTGCAAGCTGCTGGAAGGCTGAGAGAACAAAGTTAAAAAATAAAAAATCTGAGTTC 1320
DB 7260 ATTGCAAGCTGCTGGAAGGCTGAGAGAACAAAGTTAAAAAATAAAAAATCTGAGTTC 7319
QY 1321 AAGAAATCTGCATTAATAGGCAATTTACAGAGTGTAAAGTTAAATTAATTAAC 1380
DB 7320 AAGAAATCTGCATTAATAGGCAATTTACAGAGTGTAAAGTTAAATTAATTAAC 7379
QY 1381 ATGGTTTATGATATGTAATTAACATTAAGGCTCAATTTCAAAAAGTGTTCAGGGAGCTTC 1440
DB 7380 ATGGTTTATGATATGTAATTAACATTAAGGCTCAATTTCAAAAAGTGTTCAGGGAGCTTC 7439
QY 1441 CTTGAGAGGTGGCAATTTGAGCAGAGAAATGATGACACAAAGAGCTAAATCTGGAAGT 1500
DB 7440 CTTGAGAGGTGGCAATTTGAGCAGAGAAATGATGACACAAAGAGCTAAATCTGGAAGT 7499
QY 1501 TTAAGGGGAAAAAGGACGTCGAAAGGCCCTGAGGCAATTAAGAAATTTGGCTGATTC 1560
DB 7500 TTAAGGGGAAAAAGGACGTCGAAAGGCCCTGAGGCAATTAAGAAATTTGGCTGATTC 7559
QY 1561 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAAGTGGGGGCAACATGAAAGTGG 1620
DB 7560 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAAGTGGGGGCAACATGAAAGTGG 7619
QY 1621 ACGCTGAGAGTGTAGGCAAGGCGCAATGCTCTGCAAGTATTTCTTGGTCAACCAACAGA 1680
DB 7620 ACGCTGAGAGTGTAGGCAAGGCGCAATGCTCTGCAAGTATTTCTTGGTCAACCAACAGA 7679
QY 1681 GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTTAGGAAGACAGAAATTTAAATCAAC 1740
DB 7680 GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTTAGGAAGACAGAAATTTAAATCAAC 7739
QY 1741 TGTATCATTAACACAGACCTTTCTCTGTATTAAGGCTCCCAAGGATCTAGAAAGACGTA 1800
DB 7740 TGTATCATTAACACAGACCTTTCTCTGTATTAAGGCTCCCAAGGATCTAGAAAGACGTA 7799
QY 1801 AGTTAAACAAGCTCTCATTTAGCAGGGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAGTT 1860
DB 7800 AGTTAAACAAGCTCTCATTTAGCAGGGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAGTT 7859
QY 1861 CAGAGTACTCCAGTCCATGCTATGAAAAGCTCCCAAAATTTGTAACAACTGACAA 1920

Db 7860 CAGAGAGTCTCCTCAGTCCATGCTGTAAGAAAGCTCCCAATTTGTAACAACCTGACAA 7919

QY 1921 ATGCAACACCTCCCAAGCTCTCCCATTTCTCTGTGTGCGCTGGGTGGGGGGTGGG 1980

Db 7920 ATGCAACACCTCCCAAGCTCTCCCATTTCTCTGTGTGCGCTGGGTGGGGGGTGGG 7979

QY 1981 TTGGAGGGGGGAAACTTTTA 2001

Db 7980 TTGGAGGGGGGAAACTTTTA 8000

RESULT 4
AF135028 11570 bp DNA linear PRI 27-JUN-2000

LOCUS Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
DEFINITION AF135028
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 11570)
AUTHORS Yousef, G.M., and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
JOURNAL J. Biol. Chem. 274 (53), 37511-37516 (1999)
MEDLINE 20076408
PUBMED 10608802

REFERENCE 2 (bases 1 to 11570)
AUTHORS Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A., and Obiezu, C.V.
TITLE The new human kallikrein gene family: implications in
carcinogenesis
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)

JOURNAL 3 (bases 1 to 11570)
MEDLINE 21121728
PUBMED 10675891

REFERENCE Yousef, G.M., Luo, L.Y., and Diamandis, E.P.
AUTHORS Direct Submision
TITLE Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada

FEATURES
SOURCE Location/Qualifiers
1. .11570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101. .2131,2210. .2293,4762. .5023,5763. .6019,
6105. .6238,11092. .11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221. .2293,4762. .5023,5763. .6019,6105. .6238,
11092. .11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
/protein_id="AAD26429.1"
/db_xref="GI:4589283"
/translation="MATARPMMVNLCAITLALHGVTEHLANDVSCDPSNTVPS
GSNDLGAAGEDARSDSSRRITNGSDCMHPQMAALLLRPNOLYCAAVLVHPQW
LTTAAHCKRYFRRLGHYSLSPIYBSGQMFQGVKSIPIHFGSHPSNDMLIKAN
RLRPTDVARIVNSHPCPSAGTCLVSGWGTYSPOVHPKVLQCLNISVLSSKRE
DAYRQIDITMFCAGDKARSDSCGDSGSPVNCNGLQGLVSGDYPICARPNRGVYT
NLCKFTKMIQSTIQANS"

ORIGIN
Query Match 100.0%; Score 2001; DB 9; Length 11570;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACGACCATCTGTGC 60
|||||

Db 6000 GACAACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACGACCATCTGTGC 6059

QY 61 CGCCTTCATCTTTCTTCACATTTCTCATTTGTTCTCTTTGAACATGACATCTCCCTAAG 120

Db 6060 CGCCTTCATCTTTCTTCACATTTCTCATTTGTTCTCTTTGAACATGACATCTCCCTAAG 6119

QY 121 TCCCTCAGAGCTGGAATATCAGGGGTGTAAGTACAGAAAAGTGCAGAGATGCTTACCCGA 180

Db 6120 TCCCTCAGAGCTGGAATATCAGGGGTGTAAGTACAGAAAAGTGCAGAGATGCTTACCCGA 6179

QY 181 GACAGATAGATGACACCATGTTCTGCGCCGGTGAACAAGCAGGTAGAGCTCTGCCAGG 240

Db 6180 GACAGATAGATGACACCATGTTCTGCGCCGGTGAACAAGCAGGTAGAGCTCTGCCAGG 6239

QY 241 TGAGACACCTCTCTTTATTACAGAGATACACCTGAGTCCCACTGGTTAACTGAGAC 300

Db 6240 TGAGACACCTCTCTTTATTACAGAGATACACCTGAGTCCCACTGGTTAACTGAGAC 6299

QY 301 GTTGCAGAAATTCAGAAATCAGCAATTGCGCAAGCAGTGAGGCCCTGTTCTCAGAGA 360

Db 6300 GTTGCAGAAATTCAGAAATCAGCAATTGCGCAAGCAGTGAGGCCCTGTTCTCAGAGA 6359

QY 361 GCTCATACCTTAGAGTATGTTGTTAGTAAATATAGCTGAGCTGCTTATGTCATTTTC 420

Db 6360 GCTCATACCTTAGAGTATGTTGTTAGTAAATATAGCTGAGCTGCTTATGTCATTTTC 6419

QY 421 CAGTTTTTATGACCAATTAAACAGGTAAACAGGTAAACAGGTAAACAGGTAAACAGGT 480

Db 6420 CAGTTTTTATGACCAATTAAACAGGTAAACAGGTAAACAGGTAAACAGGTAAACAGGT 6479

QY 481 GTAATCCACACCTTTGGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAAGAGTTTGA 540

Db 6480 GTAATCCACACCTTTGGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAAGAGTTTGA 6539

QY 541 CTAGCCTGAGCAATGAGGCAAACTCTGCTCTTAAACAAATATACAAATATAGCTCTGC 600

Db 6540 CTAGCCTGAGCAATGAGGCAAACTCTGCTCTTAAACAAATATAGCTCTGC 6539

QY 601 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACACAGATATCACTTAA 660

Db 6600 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACACAGATATCACTTAA 6659

QY 661 ACCCAGAGGTGAGGTGAGTGAAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 720

Db 6660 ACCCAGAGGTGAGGTGAGTGAAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 7219

QY 721 AGAGTGACATTTTGTCTCAAAAGAAAAAACAATGATCAATGATGATGATGATGATGATG 780

Db 6720 AGAGTGACATTTTGTCTCAAAAGAAAAAACAATGATCAATGATGATGATGATGATGATG 6779

QY 781 TTAACCTTAATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840

Db 6780 TTAACCTTAATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6839

QY 841 ACAATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900

Db 6840 ACAATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6899

QY 901 GAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960

Db 6900 GAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6959

QY 961 AGTAGCCACATGCTGAGCAAGTAACTGATGATGATGATGATGATGATGATGATGATG 1020

Db 6960 AGTAGCCACATGCTGAGCAAGTAACTGATGATGATGATGATGATGATGATGATGATG 7019

QY 1021 AGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080

Db 7020 AGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7079

QY 1081 GGCACCTCGTGTACAGATGTCAGTTTGGCAGTTTGGCAGCTGTGTGTAGTTAGTGTCT 1140

Db 7080 GGCACCTCGTGTACAGATGTCAGTTTGGCAGTTTGGCAGCTGTGTGTAGTTAGTGTCT 7139

QY 1141 TTGTTTCACAAATCTGTAATATGACAGTTTCTAGCAAGTCTGTGTAATAATCTCTGA 1200
| | | | |
Db 7140 TTGTTTCACAAATCTGTAATATGACAGTTTCTAGCAAGTCTGTGTAATAATCTCTGA 7199
| | | | |
QY 1201 GGAAGGAAAAGAAAATCTGTGTAGTATTTTTCACAGAGATATTTAATACAGGGAGTTA 1260
| | | | |
Db 7200 GGAAGGAAAAGAAAATCTGTGTAGTATTTTTCACAGAGATATTTAATACAGGGAGTTA 7259
| | | | |
QY 1261 ATTGCAAGGCTGCTGGAAGGGCTGAGAGAACAAAGTTAAATAAATCTGTGTGTC 1320
| | | | |
Db 7260 ATTGCAAGGCTGCTGGAAGGGCTGAGAGAACAAAGTTAAATAAATCTGTGTGTC 7319
| | | | |
QY 1321 AAGATCTGCATTAATATGAGGCAATTTTCAGAGAGTGTAAAGTTAAATCCCAATATAAAC 1380
| | | | |
Db 7320 AAGATCTGCATTAATATGAGGCAATTTTCAGAGAGTGTAAAGTTAAATCCCAATATAAAC 7379
| | | | |
QY 1381 ATGGTTTTCAGATATGTAACATATAGGCGCAATATTCATAAAGGTGTCTAGGGAGCTTC 1440
| | | | |
Db 7380 ATGGTTTTCAGATATGTAACATATAGGCGCAATATTCATAAAGGTGTCTAGGGAGCTTC 7439
| | | | |
QY 1441 CTGAGAGAGGTGGCATTGAGCAGAGATGAGTACCAAGAACTAACTGTGAGT 1500
| | | | |
Db 7440 CTGAGAGAGGTGGCATTGAGCAGAGATGAGTACCAAGAACTAACTGTGAGT 7499
| | | | |
QY 1501 TTAAGGGGAAAAGAAAAGCAGTGCAGAAAGGCCCTGAGGAGTAAAGAAATTTGGCTGATTC 1560
| | | | |
Db 7500 TTAAGGGGAAAAGAAAAGCAGTGCAGAAAGGCCCTGAGGAGTAAAGAAATTTGGCTGATTC 7559
| | | | |
QY 1561 AAAAGAGAGAGGAAAACCAATGCACTGAGAAACAAAAGTGGGGGCAACAGTAAAGT 1620
| | | | |
Db 7560 AAAAGAGAGAGGAAAACCAATGCACTGAGAAACAAAAGTGGGGGCAACAGTAAAGT 7619
| | | | |
QY 1621 ACGCTGAGAGTGTAGGAGGGGGGGAATGCTGTCAAGTATTTCTGTGTCACCAACAG 1680
| | | | |
Db 7620 ACGCTGAGAGTGTAGGAGGGGGGGAATGCTGTCAAGTATTTCTGTGTCACCAACAG 7679
| | | | |
QY 1681 GCTTCCCTATGTTCTAATGGAAGCTGTATCTGTGAGAGAGACAGATTTAAATCAAC 1740
| | | | |
Db 7680 GCTTCCCTATGTTCTAATGGAAGCTGTATCTGTGAGAGAGACAGATTTAAATCAAC 7739
| | | | |
QY 1741 TGTTCATCTCAACCAAGCCTCTCTGTATTCAGGCTCCCAAGGAGCTTAAGAGAGT 1800
| | | | |
Db 7740 TGTTCATCTCAACCAAGCCTCTCTGTATTCAGGCTCCCAAGGAGCTTAAGAGAGT 7799
| | | | |
QY 1801 AGTTAACAGCTCTCATAGAGGTGTGTGTTTCAACAGATGTTAGAGAGTGGGAGT 1860
| | | | |
Db 7800 AGTTAACAGCTCTCATAGAGGTGTGTGTTTCAACAGATGTTAGAGAGTGGGAGT 7859
| | | | |
QY 1861 CAGAGAGTCTCAGTCCCATGCTATGAAAAAGCTCCCAAAATTTGTAACAACCTGCACA 1920
| | | | |
Db 7860 CAGAGAGTCTCAGTCCCATGCTATGAAAAAGCTCCCAAAATTTGTAACAACCTGCACA 7919
| | | | |
QY 1921 ATGCAACACTCCCAAGCTCTCCCATTTCTTCTGTGTGCTGTGAGTGTGGGGAGTGG 1980
| | | | |
Db 7920 ATGCAACACTCCCAAGCTCTCCCATTTCTTCTGTGTGCTGTGAGTGTGGGGAGTGG 7979
| | | | |
QY 1981 TTGCGAGGGGGAACCTTTA 2001
| | | | |
Db 7980 TTGCGAGGGGGAACCTTTA 8000
| | | | |

RESULT 5
AC011483/c 107487 bp DNA linear PRI 01-JUL-2002
LOCUS AC011483 Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
DEFINITION AC011483
AC011483.7 GI:21637461
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submision
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40.99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1..107487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C22"
misc_feature 64998..65494
/note="NOTE: Shatter libraries failed to resolve dinucleotide repeat. Unsure number of repeat copies 64998-65494. Forced join 65015."
ORIGIN
Query Match 100.0%; Score 2001; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACACAGAGGCCCAAGAGGAGTGTCCAGGTTCTTTGATACCAACCCATCTGCG 60
| | | | |
Db 54220 GACACACAGAGGCCCAAGAGGAGTGTCCAGGTTCTTTATACCAACCCATCTGCG 54161
| | | | |
QY 61 CGCTTCATCTTCTCCCACTCTCATGTTGTCTGTTTGAACAGTCCCTTAAG 120
| | | | |
Db 54160 CGCTTCATCTTCTCCCACTCTCATGTTGTCTGTTTGAACAGTCCCTTAAG 54101
| | | | |
QY 121 TCCTCAGTGTGAATATCAGCGTGTAAAGTCAGAAAAGTGCAGAGTCTTACCGA 180
| | | | |
Db 54100 TCCTCAGTGTGAATATCAGCGTGTAAAGTCAGAAAAGTGCAGAGTCTTACCGA 54041
| | | | |
QY 181 GACAGTATATGACACCAATGTTCTGCGCGGTGACAAACAGTATAGACTCTGCAAG 240
| | | | |
Db 54040 GACAGTATATGACACCAATGTTCTGCGCGGTGACAAACAGTATAGACTCTGCAAG 53981
| | | | |
QY 241 TGAGACACCTCTCTTATTCAGAGATACACATGAGCCAACTCGTAACATGAGC 300
| | | | |
Db 53980 TGAGACACCTCTCTTATTCAGAGATACACATGAGCCAACTCGTAACATGAGC 53921
| | | | |
QY 301 GTTGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGA 360
| | | | |
Db 53920 GTTGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGA 53861
| | | | |
QY 361 GCTCATACCTTGAAGTGTGTGTTTATGAGAAATATCTGAGTGTCTTATGCTATTTC 420
| | | | |
Db 53860 GCTCATACCTTGAAGTGTGTGTTTATGAGAAATATCTGAGTGTCTTATGCTATTTC 53801
| | | | |
QY 421 CAGTTTTTATGAGCACAATTAAACAGGTAAATAAGGTGGGCGCAGTGGCTCACACCT 480
| | | | |
Db 53800 CAGTTTTTATGAGCACAATTAAACAGGTAAATAAGGTGGGCGCAGTGGCTCACACCT 53741
| | | | |
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGATCACTTTGTGTCAGAGTTTGAGA 540

Db 53740 GTATCCAGACCTTTGGAGGCTGAGGAGGAGATCACTTTGGTCAGAGTTTGAGA 53681
 Qy 541 CTAGCTGGCCCAATGCGGAACTGTCTCTTAAAAAAAATCAAAAATTTAGCCCTGGC 600
 Db 53680 CTAGCTGGCCCAATGCGGAACTGTCTCTTAAAAAAAATCAAAAATTTAGCCCTGGC 53621
 Qy 601 ATGTGGCGGGCGCTGTAACTCAGCTGCTCAGAGGCGCGAGCACAGAAATCACTTAA 660
 Db 53620 ATGTGGCGGGCGCTGTAACTCAGCTGCTCAGAGGCGCGAGCACAGAAATCACTTAA 53561
 Qy 661 ACCGAGAGGTGAGGTTCAGTACGATGATGCTGCACCTCACTCCAGCTGGAGAC 720
 Db 53560 ACCGAGAGGTGAGGTTCAGTACGATGATGCTGCACCTCACTCCAGCTGGAGAC 53501
 Qy 721 AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAG 780
 Db 53500 AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAG 53441
 Qy 781 TTAACTTAATTAACCAATGTATCCAAATCAATCAATTTCAAGGTAAATTAATTA 840
 Db 53440 TTAACTTAATTAACCAATGTATCCAAATCAATCAATTTCAAGGTAAATTAATTA 53381
 Qy 841 ACAATTATGATGATGATCTTACATTTCTTTCTTTTCAATTAAGTCTTGAAGT 900
 Db 53380 ACAATTATGATGATGATCTTACATTTCTTTCTTTTCAATTAAGTCTTGAAGT 53321
 Qy 901 GAGTATATATGTATGTCTGACAGACATCTCAATTTTGACTGATCACTTCACTGCTC 960
 Db 53320 GAGTATATATGTATGTCTGACAGACATCTCAATTTTGACTGATCACTTCACTGCTC 53261
 Qy 961 AGTAGCCACATGTCGCTAGCAGTAACTGTATGTGATGGCAGGATCTTAGAGGAAAGATC 1020
 Db 53260 AGTAGCCACATGTCGCTAGCAGTAACTGTATGTGATGGCAGGATCTTAGAGGAAAGATC 53201
 Qy 1021 AGGCTGTTTTTGTATGTGGGCGAGTTGTCATCTGATTAAGATCAATATCAATAGG 1080
 Db 53200 AGGCTGTTTTTGTATGTGGGCGAGTTGTCATCTGATTAAGATCAATATCAATAGG 53141
 Qy 1081 GGCATCTCGTGTACAGATGTCACTTTTGGCAGTTTCAAGCGTGTGTAGTTAAGTGTCTC 1140
 Db 53140 GGCATCTCGTGTACAGATGTCACTTTTGGCAGTTTCAAGCGTGTGTAGTTAAGTGTCTC 53081
 Qy 1141 TTGTTTTAAACAAATCTGTATATGATGACATTTTCTACCAAGTGTGTAAATATCTTGA 1200
 Db 53080 TTGTTTTAAACAAATCTGTATATGATGACATTTTCTACCAAGTGTGTAAATATCTTGA 53021
 Qy 1201 GGAAGAAAAGAGAAATCTGGTAGTATTTTAAACAAGAAATATTTAATACAGGGAGTTA 1260
 Db 53020 GGAAGAAAAGAGAAATCTGGTAGTATTTTAAACAAGAAATATTTAATACAGGGAGTTA 52961
 Qy 1261 ATTGCAAAAGCTGCTGAGAGGCTGAGAGAAACAAAGTTAAAAAATAAAAAATCTGTGTCTC 1320
 Db 52960 ATTGCAAAAGCTGCTGAGAGGCTGAGAGAAACAAAGTTAAAAAATAAAAAATCTGTGTCTC 52901
 Qy 1321 AAGAAATCTGATTAATAGGCGCAATTTCAAGAGAGTGTAAAGTTAAACCCCAAAATAAAC 1380
 Db 52900 AAGAAATCTGATTAATAGGCGCAATTTCAAGAGAGTGTAAAGTTAAACCCCAAAATAAAC 52841
 Qy 1381 ATGGTTTAAAGATAGTAAACAAATAAAGGCGCAATTTCAAAAGAGTGTGACAGGGAGCTC 1440
 Db 52840 ATGGTTTAAAGATAGTAAACAAATAAAGGCGCAATTTCAAAAGAGTGTGACAGGGAGCTC 52781
 Qy 1441 CTGGAGAGGTGCGATTTAGCAGAGATGATGACAAAGAAAGCTTAACTCGTGAAGT 1500
 Db 52780 CTGGAGAGGTGCGATTTAGCAGAGATGATGACAAAGAAAGCTTAACTCGTGAAGT 52721
 Qy 1501 TTAAGGGGAAAAGAAAGCACTGTCAAAAGGCGCTGAGGCACTAAGAAATTTGGCTGATTC 1560
 Db 52720 TTAAGGGGAAAAGAAAGCACTGTCAAAAGGCGCTGAGGCACTAAGAAATTTGGCTGATTC 52661
 Qy 1561 AAAGAAAGAGAGAAACCAATGCAACTGAGAAACAAAGAGTGGGGCAACAGTAAAGTGT 1620

Db 52660 AAAGAAAGAGAGAAACCAATGCACTGAGAAACAAAGTGGGGCAACAGTAAAGTGTG 52601
 Qy 1621 ACCGTGAGAGTGTAGGCAAGGGGGAATGCTCTCAAGTATTTCTGTCTACCAACAGCA 1680
 Db 52600 ACCGTGAGAGTGTAGGCAAGGGGGAATGCTCTCAAGTATTTCTGTCTACCAACAGCA 52541
 Qy 1681 GCTTCCCTATGTTCTAATAGGAGCTGTATCTGTGTGAGAGAGACAAATTTAAATCAAC 1740
 Db 52540 GCTTCCCTATGTTCTAATAGGAGCTGTATCTGTGTGAGAGAGACAAATTTAAATCAAC 52481
 Qy 1741 TGTTCATCAACAGCAACCCCTTCTGTATTCAGGCTCCCAAGGATCTAAGAGACGTA 1800
 Db 52480 TGTTCATCAACAGCAACCCCTTCTGTATTCAGGCTCCCAAGGATCTAAGAGACGTA 52421
 Qy 1801 AGTTAACAGCTCTCATTTAGCAGGCTGTGTTCACACAGTATTAAGAACTGGGGATT 1860
 Db 52420 AGTTAACAGCTCTCATTTAGCAGGCTGTGTTCACACAGTATTAAGAACTGGGGATT 52361
 Qy 1861 CAGAGTACTCCAGTCCCAATGCTATGAAAAGCTCCCAATTTGTACAAACCTGACAA 1920
 Db 52360 CAGAGTACTCCAGTCCCAATGCTATGAAAAGCTCCCAATTTGTACAAACCTGACAA 52301
 Qy 1921 ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCTGGGTGGGGGGTGGG 1980
 Db 52300 ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCTGGGTGGGGGGTGGG 52241
 Qy 1981 TTGCGAGGGGGAACCTTTTA 2001
 Db 52240 TTGCGAGGGGGAACCTTTTA 52220

RESULT 6
 AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
 LOCUS AF243527 Homo sapiens serine protease gene cluster, complete sequence.
 DEFINITION AF243527
 ACCESSION AF243527.1 GI:11244757
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 Gen, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McQuaig, J.,
 Moss, P., Paepel, B. and Wang, K.
 1 (bases 1 to 230000)
 TITLE
 JOURNAL
 Gene 257 (1), 119-130 (2000)
 MEDLINE
 20510030
 PUBMED
 11054574
 REFERENCE
 2 (bases 1 to 230000)
 Gen, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McQuaig, J.,
 Moss, P., Paepel, B. and Wang, K.
 Direct Submission
 TITLE
 Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
 Bothell, WA 98021, USA
 FEATURES
 source
 1. 230000
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q13"
 /complement(<6963..>11517)
 /gene="KLK1"
 complement(join(<6963..7118,7668..7804,7923..8212,
 9482..9641,11472..11517))
 /gene="KLK1"
 /product="renal kallikrein"
 complement(join(6963..7118,7668..7804,7923..8212,
 9482..9641,11472..11517))
 /gene="KLK1"
 /note="serine protease"

/codon_start=1
 /product="renal kallikrein"
 /protein_id="AA033353.1"
 /db_xref="GI:11244758"
 /translation="MFWPLVCLALSLGTGAAPPIOSRIIVGSECEHQSOPWQALYH
 FSTFOCGGILLVHQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFN
 MSLENNHROADEYDSHDMILRLTEPADTITAVKVELPTEPEVSGTCLASGNS
 IEPNESPDDIQCVDLKITLPNDCCKAHOKXTDPMFLCYGHLGGKDKTCVDSGGPL
 MCDGLVGTSMGIVPCGTPKPEVANRVLSYKMLIEDTIAENS"
 complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..>19246))
 /product="ACO protease"
 /complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..19246))
 /note="serine protease"
 /codon_start=1
 /product="ACO protease"
 /protein_id="AA033354.1"
 /db_xref="GI:11244759"
 /translation="MMLVLTLSFLTAATAODGDKLEGBDECAHSPOMQVALYGR
 FNCASLISPPWVLSAHCOSRPMRVVLGSHNLKXDGPEQLATTSRVLPDRYERS
 HRNDIMILRIYQPARLNPQVRPALPLPRCHPBEACVCSMGKIVLSSVSPLDTLHC
 ANITISDTSCKSPGLTNVIVYCAABRSABSCBDSGGPLVCGGILQGISWGD
 VPCNTKPGYIVYCVHLEWIRETKRN"
 <42595..>47769
 /gene="KLK3"
 join(<42595..42640,43880..44039,45669..45955,46099..46235,
 47614..>47769)
 /gene="KLK3"
 /product="prostate specific antigen"
 join(42595..42640,43880..44039,45669..45955,46099..46235,
 47614..47769)
 /gene="KLK3"
 /note="serine protease"
 /codon_start=1
 /product="prostate specific antigen"
 /protein_id="AA033355.1"
 /db_xref="GI:11244760"
 /translation="MWWPVVFLTSLVTWIGAADPLLSRIIVGSECEHQSOPWQVAVS
 RGRVACGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFLYD
 MSELNRLRFGDDSHDMILRLSEPAELTDVKKWMLDLPORPALGTTCVASGMSI
 EPEELRPKIKCYVDLIVINDVCAQNPQKTKKMLCGRWTKGKSTCSGDSGGPLV
 CNGVLQGTSMGSEPCALPERPISLYTKVHTRKMIKDTIYANP"
 <61139..>66229
 /gene="KLK2"
 join(<61139..61184,62391..62550,64142..64428,64542..64678,
 66074..>66229)
 /gene="KLK2"
 /product="glandular kallikrein"
 join(61139..61184,62391..62550,64142..64428,64542..64678,
 66074..66229)
 /gene="KLK2"
 /note="serine protease"
 /codon_start=1
 /product="glandular kallikrein"
 /protein_id="AA033356.1"
 /db_xref="GI:11244761"
 /translation="MMDLVLSIALSVGTGAVPLIOSRIIVGSECEHQSOPWQVAVS
 HGMHACGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFLYN
 MSLKHSRLEDEDSHDMILRLSEPAKITDVKKVLGILDTORPALGTTCVASGMSI
 EPEELRPRSICQVSHLINDMCAARVSEKIVKPMICGLMTGKADTCTGDSGGGLV
 CNGVLQGTSMGSEPCALPERPVTYTKVHTRKMIKDTIYANP"
 complement(join(94602..94754,96027..96163,96247..96497,
 96919..97081,98345..>98405))
 /product="serine protease"
 /complement(join(94602..94754,96027..96163,96247..96497,
 96919..97081,98345..98405))
 /codon_start=1
 /product="serine protease"
 /protein_id="AA033357.1"
 /db_xref="GI:11244762"
 /translation="MATAQPMWVFLGYTLIVAGSLVSGSCSQIINDGDCSPHSPW
 QALVMEHELCSGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFN
 MSLENNHROADEYDSHDMILRLTEPADTITAVKVELPTEPEVSGTCLASGNS
 IEPNESPDDIQCVDLKITLPNDCCKAHOKXTDPMFLCYGHLGGKDKTCVDSGGPL
 MCDGLVGTSMGIVPCGTPKPEVANRVLSYKMLIEDTIAENS"
 complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..>19246))
 /product="ACO protease"
 /complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..19246))
 /note="serine protease"
 /codon_start=1
 /product="ACO protease"
 /protein_id="AA033358.1"
 /db_xref="GI:11244763"
 /translation="MMLVLTLSFLTAATAODGDKLEGBDECAHSPOMQVALYGR
 FNCASLISPPWVLSAHCOSRPMRVVLGSHNLKXDGPEQLATTSRVLPDRYERS
 HRNDIMILRIYQPARLNPQVRPALPLPRCHPBEACVCSMGKIVLSSVSPLDTLHC
 ANITISDTSCKSPGLTNVIVYCAABRSABSCBDSGGPLVCGGILQGISWGD
 VPCNTKPGYIVYCVHLEWIRETKRN"
 <42595..>47769
 /gene="KLK3"
 join(<42595..42640,43880..44039,45669..45955,46099..46235,
 47614..>47769)
 /gene="KLK3"
 /product="prostate specific antigen"
 join(42595..42640,43880..44039,45669..45955,46099..46235,
 47614..47769)
 /gene="KLK3"
 /note="serine protease"
 /codon_start=1
 /product="prostate specific antigen"
 /protein_id="AA033359.1"
 /db_xref="GI:11244764"
 /translation="MWWPVVFLTSLVTWIGAADPLLSRIIVGSECEHQSOPWQVAVS
 RGRVACGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFLYD
 MSELNRLRFGDDSHDMILRLSEPAELTDVKKWMLDLPORPALGTTCVASGMSI
 EPEELRPKIKCYVDLIVINDVCAQNPQKTKKMLCGRWTKGKSTCSGDSGGPLV
 CNGVLQGTSMGSEPCALPERPISLYTKVHTRKMIKDTIYANP"
 <61139..>66229
 /gene="KLK2"
 join(<61139..61184,62391..62550,64142..64428,64542..64678,
 66074..>66229)
 /gene="KLK2"
 /product="glandular kallikrein"
 join(61139..61184,62391..62550,64142..64428,64542..64678,
 66074..66229)
 /gene="KLK2"
 /note="serine protease"
 /codon_start=1
 /product="glandular kallikrein"
 /protein_id="AA033360.1"
 /db_xref="GI:11244765"
 /translation="MMDLVLSIALSVGTGAVPLIOSRIIVGSECEHQSOPWQVAVS
 HGMHACGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFLYN
 MSLKHSRLEDEDSHDMILRLSEPAKITDVKKVLGILDTORPALGTTCVASGMSI
 EPEELRPRSICQVSHLINDMCAARVSEKIVKPMICGLMTGKADTCTGDSGGGLV
 CNGVLQGTSMGSEPCALPERPVTYTKVHTRKMIKDTIYANP"
 complement(join(94602..94754,96027..96163,96247..96497,
 96919..97081,98345..>98405))
 /product="serine protease"
 /complement(join(94602..94754,96027..96163,96247..96497,
 96919..97081,98345..98405))
 /codon_start=1
 /product="serine protease"
 /protein_id="AA033361.1"
 /db_xref="GI:11244766"
 /translation="MATAQPMWVFLGYTLIVAGSLVSGSCSQIINDGDCSPHSPW
 QALVMEHELCSGVLVHPQWVLTAAHCISDNQMLMGRHNLFDENLTQFVHSESPHPPFN
 MSLENNHROADEYDSHDMILRLTEPADTITAVKVELPTEPEVSGTCLASGNS
 IEPNESPDDIQCVDLKITLPNDCCKAHOKXTDPMFLCYGHLGGKDKTCVDSGGPL
 MCDGLVGTSMGIVPCGTPKPEVANRVLSYKMLIEDTIAENS"
 complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..>19246))
 /product="ACO protease"
 /complement(join(13552..13704,14377..14530,14678..14917,
 15416..15566,19204..19246))
 /note="serine protease"
 /codon_start=1
 /product="ACO protease"
 /protein_id="AA033362.1"
 /db_xref="GI:11244767"
 /translation="MMLVLTLSFLTAATAODGDKLEGBDECAHSPOMQVALYGR
 FNCASLISPPWVLSAHCOSRPMRVVLGSHNLKXDGPEQLATTSRVLPDRYERS
 HRNDIMILRIYQPARLNPQVRPALPLPRCHPBEACVCSMGKIVLSSVSPLDTLHC
 ANITISDTSCKSPGLTNVIVYCAABRSABSCBDSGGPLVCGGILQGISWGD
 VPCNTKPGYIVYCVHLEWIRETKRN"
 <42595..>47769
 /gene="KLK3"
 join(<42595..42640,43880..44039,45669..45955,46099..46235,
 47614..>

VRHPEYNNRLIANDLMIKIDUESSESDTRISIAQCCPPLAGNSCLVSGMGLIANR
MPTYLICAKPVCSVEBVCYSKLVDLYNHPSMFCAGGHDQKDSCNBDSGGFLINCGLYIQ
GLTVFACAPCGQVGPVTYNLCKFETMEITXCTVAS"
gene complement (<131301..>140303)
/gene="CLK5"
/note="synonym: SCTE"
mRNA complement {join(<131301..131456,136310)..136443,
136529..136785,137525..137690,140255..>140303)}
/gene="CLK5"
/product="stratum corneum trypsin-like serine protease"
CDS complement {join(131301..131456,136310..136443,
136529..136785,137525..137690,140255..140303)}
/gene="CLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MMWVLCALITALTLGVTDARSDSSRIINSDCDMTWPQVAL
ILRPALTCGAIVLHPOMLLTAHCRKKVFVIRLGHSLSIPVESQGOMFGVKSIFH
PGYSHPGNSNDLMILIKLNRRIRTPKVDPVIRPVSCHSAGKCLVSGMGTTSPQVHF
PKVLGINSTVLSQRCEADAPROIDDTPMPAGDKAGRDSCQDSDSGSPVVCNSSLGSL
VSKGDYPCARNRBRGYTNLCKFETKMIOERTIQANS"
mRNA complement {join(<146834..146966,149628..149764,
151186..151433,155052..155208,155948..>155987)}
/product="protease M"
CDS complement {join(146834..146966,149628..149764,
151186..151433,155052..155208,155948..155987)}
/note="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKKLNVVLSLIAAMAEEQNKLHVGPCDTSHPYOALYTSGH
LLCGSVLIHLPLMWLTAAHCCKKNLOVFEIGKNLRISQSOESGVADVHTPIPDAS
HDODILHLPMLTAAHCRKKPNLOVFEIGKNLRISQSOESGVADVHTPIPDAS
HLVSEEBEENHYPCQITTONMTCADDEKYGKDCSCGDSBGFLVCGDHLRGLVSGNIIPC
GSKEKPYGTIVNCRYTWMIQKTIOAK"
gene complement (<165420..>170283)
/gene="CLK7"
/note="synonym: SCCE"
mRNA complement {join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283)}
/gene="CLK7"
/product="stratum corneum chymotryptic enzyme"
CDS complement {join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..170283)}
/gene="CLK7"
/note="serine protease"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLILPLQILLSIALLETAGEAQGDKIIIDAPARGSHGPHQ
VALLSGNDLHGCVLVNERWVLTAAHKCMNEVTHLASDTLDGDRRAQIRASKSFRRP
GYSQTUHNDMILVKINSOARLSISMWKVYLPSCRCPPGTTCVSGMGTTSPDPYPR
SDLMCVDVYALKIPDDCTKYTKDLLENMSLCAIGIDPSKKNAANGSGGPLVCRGLQSL
VSWGFPPCGQNDSDGVTYQVCKDFETKMINDTKKRH"
mRNA complement {join(<183943..184098,185635..185768,
187865..188127,188293..188452,188957..>189036)}
/product="neutropin"
CDS complement {join(183943..184098,185635..185768,
187865..188127,188293..188452,188957..189036)}
/note="serine protease; also called ovaasin"
/codon_start=1
/product="neutropin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRRPRPAATATMMFTLLIGAMVGHSAEDKVLGGHEOCPSHQ
PMQALTFGQQILLCGVTVAGSNWVLTAAHCRKKPYTRYRLDHSILANQDSRPORTPVVO
SIHPHCNTSVDEBNHMDLQLORDASLDSJCTPISLADHCKVLSGCTVSGEWTV
TSPRNPFSDTDAENCAEVKIFPOKCEDAVSQGITDGWCAAGSKPADPCQDSDGGPLVVC
DGALEGISWMSDSCGRSDKGVTYTNCRILYDIWKIKLIGSGH"
gene complement {join(<190980..191129,191573..191709,
192011..192144,192597..192730,193081..193214,193565..>193699)}

CDS

194324..194589,197048..197204,193730..>197412))
/product="kallikrein-like 3"
/complement(join(190980..191129,191573..191709,
194324..194589,197048..197204,193730..197412))
/note="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAC33362.1"
/db_xref="gi:11244767"
/translation="MKUGLCLALSLSLAGHMDTRAIAGEECPNSOPMQAGLHLT
RLFGATLISDRMLTAHCKRKYLWRLKGEHMLKGEPLRVDFPHGFND
LSANDHDIWLIRLPDARLSPAPVPLNSQTSVGMASVSPKLLFV
TLQCANISILENKLCHNAYPGHISDSMLCAGLMEGSGSCGDSGGLVNCGLTAGVY
SGAEPCSRPRRAVYTSVCHYLDIMQIEIMEN"

Query Match 100.0%; Score 2001; DB 9; Length 230000;

Beeb Local Similarity 100.0%; Pred. No. 0;

Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAACAGAGCCCAAGGTGAGTCAGCTTCTTGATACCGAACCATCTCTGC 60
DB 136548 GAGAACCAAGAGCCCCCAAGGTGAGTCAGCTTCTTGATACCGAACCATCTCTGC 136489
QY 61 CGCCTTCATCTTCTCCACTTCTCAATGTGTCCTGTTGACAGTGACCTTCCCTAAG 120
DB 136488 CGCTTCATCTTCTCCACTTCTCAATGTGTCCTGTTGACAGTGACCTTCCCTAAG 136429
QY 121 TCTCCAGTCTTGAATATCAGCGTGCTAAGTCAAGAAAGGTGCGAGAGTCTTACCGGA 180
DB 136428 TCTCCAGTCTTGAATATCAGCGTGCTAAGTCAAGAAAGGTGCGAGAGTCTTACCGGA 136369
QY 181 GACAGATGATGACCAATGTTTCCGCGGTGACAAAGCGATGAGACCTCTGCGCAAG 240
DB 136368 GACAGATGATGACCAATGTTTCCGCGGTGACAAAGCGATGAGACCTCTGCGCAAG 136309
QY 241 TGAGAGACCTCTCTTATTCAGAGATACACACTGAGTCCAACTCGGTAACTGAGAGC 300
DB 136308 TGAGAGACCTCTCTTATTCAGAGATACACACTGAGTCCAACTCGGTAACTGAGAGC 136249
QY 301 GTTGCCAAATTCGAGAAATCCAGAAATTCGCAAGACAGTCCGCTGTTCTACAGA 360
DB 136248 GTTGCCAAATTCGAGAAATCCAGAAATTCGCAAGACAGTCCGCTGTTCTACAGA 136189
QY 361 GCTCATACCTAGAGTATGCTTATAGTAATAATGCTGAGCTGTTATGTCATTTTC 420
DB 136188 GCTCATACCTAGAGTATGCTTATAGTAATAATGCTGAGCTGTTATGTCATTTTC 136129
QY 421 CAGTTTTTATAGTACCAATTAAGAAAGGTAAAGAGGTGGGCGAGTGGCTCAACCT 480
DB 136128 CAGTTTTTATAGTACCAATTAAGAAAGGTAAAGAGGTGGGCGAGTGGCTCAACCT 136069
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGACGAGATCACTTTTGTCAGAGAGTTTGA 540
DB 136068 GTAATCCAGCACTTTGGAGGCTGAGGACGAGATCACTTTTGTCAGAGAGTTTGA 136009
QY 541 CTACCTGGCCAAATGCGGAAATCTGCTCTTAAAAAAATCAAAAAATTACCTGGC 600
DB 136008 CTACCTGGCCAAATGCGGAAATCTGCTCTTAAAAAAATCAAAAAATTACCTGGC 135949
QY 601 ATGTGCGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCACTTAA 660
DB 135948 ATGTGCGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCACTTAA 135889
QY 661 ACCCAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCAACTGGAGAC 720
DB 135888 ACCCAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCAACTGGAGAC 135829
QY 721 AGAGTGCACCTTTGCTCAAAAAGAAAAAAACAGTAAAAAAAGAAACAGGTGAAG 780
DB 135828 AGAGTGCACCTTTGCTCAAAAAGAAAAAAACAGTAAAAAAAGAAACAGGTGAAG 135769
QY 781 TTAACCTTAAATCAATGATCCCAATCAATCAATCAATCAATCAATCAATCAATCAAT 840

DB 135768 TTAACCTTAAATCAATGATCCCAATCAATCAATCAATCAATCAATCAATCAAT 135709
QY 841 ACAATTTGAATGAGATACCTTTACATTTCTTTCTGTTTCAATTAAGCTTTGAAGT 900
DB 135708 ACAATTTGAATGAGATACCTTTACATTTCTTTCTGTTTCAATTAAGCTTTGAAGT 135649
QY 901 GAGATATATGTTATGCTGACAGACATCTCAATTTTGAATGATCACTTACAGTGTCTC 960
DB 135648 GAGATATATGTTATGCTGACAGACATCTCAATTTTGAATGATCACTTACAGTGTCTC 135589
QY 961 AGTAGCCACATGTGGCTAGCAGTTACTGTATTTGATGACCGAGATCTAGAGGAAAGATC 1020
DB 135588 AGTAGCCACATGTGGCTAGCAGTTACTGTATTTGATGACCGAGATCTAGAGGAAAGATC 135529
QY 1021 AGGCTGTTTTGTTATGCTTGGGCGAGTGTGCACTGATTAAGATATCAATATAGG 1080
DB 135528 AGGCTGTTTTGTTATGCTTGGGCGAGTGTGCACTGATTAAGATATCAATATAGG 135469
QY 1081 GGCACCTCGTTATACAGATGTCAGTTTGGCAGTTTCAAGCGGTGTAGTAAAGTGC 1140
DB 135468 GGCACCTCGTTATACAGATGTCAGTTTGGCAGTTTCAAGCGGTGTAGTAAAGTGC 135409
QY 1141 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGTGTAAATATCTTGA 1200
DB 135408 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGTGTAAATATCTTGA 135349
QY 1201 GGAAGAAAAGAAATCTGTATGTTTATTTTCAAGAGATTTTAAATCAGGGGATTA 1260
DB 135348 GGAAGAAAAGAAATCTGTATGTTTATTTTCAAGAGATTTTAAATCAGGGGATTA 135289
QY 1261 ATTGCAAGCTGTGGAAGGGCTGGAAGAACAAAGTTAAATTAATTAATTAATTAATTA 1320
DB 135288 ATTGCAAGCTGTGGAAGGGCTGGAAGAACAAAGTTAAATTAATTAATTAATTAATTA 135229
QY 1321 AAGATCTGATTAATAGGCAATTTCAAGAGTGTGTAAGTTAAATCCCAAAATTAATTAAC 1380
DB 135228 AAGATCTGATTAATAGGCAATTTCAAGAGTGTGTAAGTTAAATCCCAAAATTAATTAAC 135169
QY 1381 ATGTTTTAGATATGTAATTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 1440
DB 135168 ATGTTTTAGATATGTAATTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 135109
QY 1441 CTTGAGAGGTGTGCAATTTGAGCAGAGATGAGTGAACAAAGAGCTAACTGTGAAGT 1500
DB 135108 CTTGAGAGGTGTGCAATTTGAGCAGAGATGAGTGAACAAAGAGCTAACTGTGAAGT 135049
QY 1501 TTAAGGGGAAAGAAAGGCACTGCAAAAGGCTGTAGAGCGTAAAGAAATTTGGCTGATTC 1560
DB 135048 TTAAGGGGAAAGAAAGGCACTGCAAAAGGCTGTAGAGCGTAAAGAAATTTGGCTGATTC 134989
QY 1561 AAGAGAGAGAGAAACCAATGCACTGAGAACAAAGTGGGGCAACGTAGAAAGT 1620
DB 134988 AAGAGAGAGAGAAACCAATGCACTGAGAACAAAGTGGGGCAACGTAGAAAGT 134929
QY 1621 ACGCTGAGGTGTAGGCAAGGGGGAATGCTCTGCAAGTATTTCTTGCTCAACCAACAGA 1680
DB 134928 ACGCTGAGGTGTAGGCAAGGGGGAATGCTCTGCAAGTATTTCTTGCTCAACCAACAGA 134869
QY 1681 GCTTCCCTATGTTCTAATGAACTGTATCTGTTGAGAGAGACAGATTTTAAATCAAC 1740
DB 134868 GCTTCCCTATGTTCTAATGAACTGTATCTGTTGAGAGAGACAGATTTTAAATCAAC 134809
QY 1741 TGTTCATCAACAGCAACCTTCTGTATTCAGCTCTCCCAAGGATCTTAGAGACGTA 1800
DB 134808 TGTTCATCAACAGCAACCTTCTGTATTCAGCTCTCCCAAGGATCTTAGAGACGTA 134749
QY 1801 AGTTAACAGCTCTCAATTTGAGAGGTTGTTTCAAGTATGTTGAGAGAGCTGGGGATT 1860
DB 134748 AGTTAACAGCTCTCAATTTGAGAGGTTGTTTCAAGTATGTTGAGAGAGCTGGGGATT 134689
QY 1861 CAGAGTACTCCAGTCCCAATGCTATGAAAGCTCCCAATTTGATCAAACTGTACAA 1920
DB 134688 CAGAGTACTCCAGTCCCAATGCTATGAAAGCTCCCAATTTGATCAAACTGTACAA 134629

QY 1921 ATGCAACACCTCCCGAGCTCTCCCATTTCTCTGTCGCTGGGNGTGGGGGGG 1980
 |||||
 Db 134628 ATGCAACACTCCCGAGCTCTCCCATTTCTCTGTCGCTGGGNGTGGGGGGG 134569
 |||||
 QY 1981 TTGCGAGGGGGAAACTTTTA 2001
 |||||
 Db 134568 TTGCGAGGGGGAAACTTTTA 134548
 |||||

RESULT 7
 AC027602/c 217346 bp DNA linear HTG 24-AUG-2002
 LOCUS Homo sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT
 DEFINITION
 AC027602
 AC027602.4 GI:11178143
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 217346)
 Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bascien,V., Beda,F., Bogunlavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehocsky,J., Levine,R., Liu,G., Liu,G., Locke,K., MacDonald,P., Margulis,N., McCarty,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 217346)
 Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bascien,V., Beda,F., Bogunlavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K., Lamazares,R., Landers,T., Lehocsky,J., Levine,R., Liu,G., Liu,G., MacDonald,P., Margulis,N., McCarty,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission

JOURNAL COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 15, 2000 this sequence version replaced gi:11136831.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9166
 Center clone name: 795 B 6
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 31% of reads
 Sequencing vector: Plasmid; n/a; 69% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 210748 bases at least Q40
 Consensus quality: 213655 bases at least Q30
 Consensus quality: 215058 bases at least Q20
 Insert size: 19400; agarose-fp
 Insert size: 216246; sum-of-contigs
 Quality coverage: 11.9 in Q20 bases; agarose-fp
 Quality coverage: 10.7 in Q20.
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1 8149: contig of 8149 bp in length
 * 8150 8249: gap of 100 bp
 * 8250 9592: contig of 1343 bp in length
 * 9593 9692: gap of 100 bp
 * 9693 10733: contig of 1041 bp in length
 * 10734 10833: gap of 100 bp
 * 10834 13519: contig of 2666 bp in length
 * 13520 13519: gap of 100 bp
 * 13620 17510: contig of 3891 bp in length
 * 17511 17610: gap of 100 bp
 * 17611 24602: contig of 6992 bp in length
 * 24603 24702: gap of 100 bp
 * 24703 35434: contig of 10732 bp in length
 * 35435 35534: gap of 100 bp
 * 35535 124474: contig of 88940 bp in length
 * 124475 124574: gap of 100 bp
 * 124575 134664: contig of 10090 bp in length
 * 134665 134764: gap of 100 bp
 * 134765 162343: contig of 27579 bp in length
 * 162344 162443: gap of 100 bp
 * 162444 208917: contig of 46674 bp in length
 * 208918 209017: gap of 100 bp
 * 209018 217346: contig of 8329 bp in length.
 Location/Qualifiers
 1. 217346
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19"
 /clone="RP11-795B6"
 /clone_1fb="RP11-11 Human Male BAC"
 1. 8149
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 misc_feature 8250..9592
 /note="assembly_fragment"
 misc_feature 9693..10733

QY 1801 AGTTAACAAGCTCTCATTTAGACAGGAGTGTGTTTCAACAGTACTAGGAAGTGGGATT 1860
DB 196613 AGTTAACAAGCTCTCATTTAGACAGGAGTGTGTTTCAACAGTACTAGGAAGTGGGATT 196554
QY 1861 CAGGAGTACTCCAGCTCCCATGCTATGAAAAAGTCCCAATTGTACAAACTGACAA 1920
DB 196553 CAGGAGTACTCCAGCTCCCATGCTATGAAAAAGTCCCAATTGTACAAACTGACAA 196494
QY 1921 ATGCAACACTCCCGAGCTCTCCCAATTCTCTGTCGCCGAGGAGGAGGAGTGGG 1980
DB 196493 ATGCAACACTCCCGAGCTCTCCCAATTCTCTGTCGCCGAGGAGGAGGAGTGGG 196434
QY 1981 TTGCGAGGGGGAACCTTTTA 2001
DB 196433 TTGCGAGGGGGAACCTTTTA 196413

RESULT 8
AC130782/c
LOCUS
DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
AC130782 200792 bp DNA linear HTG 22-NOV-2002
ordered pieces.
AC130782
AC130782.2 GI:25167101
VERSION HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE
AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carigala, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Kari, P., Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B., Margulis, E.H., Mastello, C., Maekari, B., McDowell, J., Paguitigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Staniripod, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 200792)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
REFERENCE 3 (bases 1 to 200792)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
On Nov 22, 2002 this sequence version replaced gi:22218452.

Center: NIH Intramural Sequencing Center
Genome Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov

Project Information
Center project name: dhz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence

contigs have been trimmed away, and each base is associated with a phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 28306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length
* 37957 73522: contig of 3566 bp in length
* 73523 73623: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88917: contig of 5150 bp in length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 3468 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193444: gap of unknown length
* 193445 199363: contig of 5919 bp in length
* 199364 199464: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

FEATURES

source

1..200792

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="CH251-355A20"

/clone_11b="CH251"

1..28306

/note="assembly_fragment"

clone end:SP6

vector side:left"

28407..37856

/note="assembly_fragment"

37957..73522

/note="assembly_fragment"

73623..83567

/note="assembly_fragment"

83668..88817

/note="assembly_fragment"

88918..125611

/note="assembly_fragment"

125712..159879

/note="assembly_fragment"

159980..174698

/note="assembly_fragment"

174799..186382

/note="assembly_fragment"

misc_feature 186483..193344
/note="assembly_fragment"
misc_feature 193445..199363
/note="assembly_fragment"
misc_feature 199464..200792
/note="assembly_fragment"
clone end:T7
vector_side:right"

ORIGIN

Query Match 96.3%; Score 1927.2; DB 2; Length 200792;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1975; Conservative 0; Mismatches 23; Indels 4; Gaps 3;
QY 1 GACACCAAGAGCCCAAGGTGAGTCCAGGTTCTTTGATACCGACCATCTGCG 60
DB 73157 GACAAACCAAGAGCCCAAGGTGAGTCCAGGTTCTTTGATACCGACCATCTGCG 73098
QY 61 CGCCTTCATCTTTCTCCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 120
DB 73097 CGCCTTCATCTTTCTCCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 73038
QY 121 TCCCTCAGTGTGTAATATCAGCGTCTAAGTCAGAAAAGGTGCGAGAGTCTTACCGA 180
DB 73037 TCCCTCAGTGTGTAATATCAGCGTCTAAGTCAGAAAAGGTGCGAGAGTCTTACCGA 72978
QY 181 GACGATAGATGACACCATGTTCTGCGCCGCTGACAAAGCAGGTAGACCTCTGCCAAG 240
DB 72977 GACGATAGATGACACCATGTTCTGCGCCGCTGACAAAGCAGGTAGACCTCTGCCAAG 72918
QY 241 TGAAGACACCTCTCTTTATTGAGAGATACACTGAGTGCCTGCGTAACTGAGAGC 300
DB 72917 TGAAGACACCTCTCTTTATTGAGAGATACACTGAGTGCCTGCGTAACTGAGAGC 72858
QY 301 GTTCCCAATCTTGAGATCCAGCAATTCAGAGACAGTCAAGACCCCTGTTCTACAGA 360
DB 72857 GTTCCCAATCTTGAGATCCAGCAATTCAGAGACAGTCAAGACCCCTGTTCTACAGA 72798
QY 361 GCTCATACCTGAGTGTGTTTGTAGTAATAATGCTGAGCTGTTATGTCATTTTC 420
DB 72797 GCTCATACCTGAGTGTGTTTGTAGTAATAATGCTGAGCTGTTATGTCATTTTC 72738
QY 421 CAGTTTTTGTAGTCCACATTTAAACAGGTAAACAGGTGCGGCGAGTGCCTACACT 480
DB 72737 CAGTTTTTGTAGTCCACATTTAAACAGGTAAACAGGTGCGGCGAGTGCCTACACT 72678
QY 481 GTATCCAGCACTTTGAGAGGTGAGGCAAGTCACTTTGTCAGAGTTTGAGA 540
DB 72677 GTATCCAGCACTTTGAGAGGTGAGGCAAGTCACTTTGTCAGAGTTTGAGA 72618
QY 541 CTAGCTGCGCCACATGCGGAACTCTGTCTTAAAAAAATACAAATTTAGCTGGC 600
DB 72617 CTAGCTGCGCCACATGCGGAACTCTGTCTTAAAAAAATTTAGCTGGC 72558
QY 601 ATGGTGGCGGCGCTGTAATCTCAGGCTCAGAGGCGGAGACAGAAATCACTTAA 660
DB 72557 ATGGTGGCGGCGCTGTAATCTCAGGCTCAGAGGCGGAGACAGAAATCACTTAA 72498
QY 661 ACCCAGAGGTGAGGTGTCAGTGAAGTGTGCACT -CACTTCAACTGGAGAGA 719
DB 72497 ACCCAGAGGTGAGGTGTCAGTGAAGTGTGCACTGCACTCAAACCTGGAGAGA 72438
QY 720 CAGAGTACACTTTTGTCTCAAAAAAGAAAAAACAGTAAAAAGAAACAGGTGAA 779
DB 72437 CAGAGTACACTTTTGTCTCAAAAAAG -AAAAAGAAAAACAGTAAAAAGAAACAGGTGAA 72379
QY 780 GTTAACCTTTAATACCAATGATCCCAATATACATTTCAAGGTGTAATTAATTA 839
DB 72378 GTTAACCTTTAATACCAATGATCCCAATATACATTTCAAGGTGTAATTAATTA 72319
QY 840 AACCAATTATGATAGATACCTTTAATCTTTCTGTTTCAATATTAAGTCTTGAAG 899
DB 72318 AACCAATTATGATAGATACCTTTAATCTTTCTGTTTCAATATTAAGTCTTGAAG 72259

QY 900 TGAATATATATGTTATGCTGACAGCATCTCAATTTGACCTAGCTACATTTGAGTGTCT 959
DB 72258 TGAATATATATGTTATGCTGACAGCATCTCAATTTGACCTAGCTACATTTGAGTGTCT 72199
QY 960 CAGTACGACATGCTGCTAGCACTTACTGATTTGATGAGTGGCAGCGATCTAGAGGAAAGAT 1019
DB 72198 CAGTACGACATGCTGCTAGCACTTACTGATTTGATGAGTGGCAGCGATCTAGAGGAAAGAT 72139
QY 1020 CAGGCTGTTTGTATGTTGAGGAGGTTGTCATGCAATTAAGATACCATATCTAATAG 1079
DB 72138 CAGGCTGTTTGTATGTTGAGGAGGTTGTCATGCAATTAAGATACCATATCTAATAG 72079
QY 1080 GGGCACTCCGTTTACAGATGATGATTTGGAGTTTTCAGGGGTGTGATTAAGTGT 1139
DB 72078 GGGCACTCCAGTTTACAGATGATGATTTGGAGTTTTCAGGGGTGTGATTAAGTGT 72019
QY 1140 CTGTTTCAACAAATCTGTAATATGACAGTTTCTAGCAAGTGTGTAATATCTTTC 1199
DB 72018 CTGTTTCAACAAATCTGTAATATGACAGTTTCTAGCAAGTGTGTAATATCTTTC 71959
QY 1200 AGAAGAAAGAAAGAAATCTGTAGGTATTTTACAGAGATATTTAATACAGGGAAT 1259
DB 71958 AGAAGAAAGAAAGAAATCTGTAGGTATTTTACAGAGATATTTAATACAGGGAAT 71899
QY 1260 AATTGCAAGCTGTGAGAGGCTGAGAGAAACAAAGTTAAATTAATTAATTAATCTGTG 1319
DB 71898 AATTGCAAGCTGTGAGAGGCTGAGAGAAACAAAGTTAAATTAATTAATTAATCTGTG 71839
QY 1320 CAAGAAATCTGATTAATAGGCAATTTACAGAGAGTGTAAAGTTAAACCCCAATTAAT 1379
DB 71838 CAAGAAATCTGATTAATAGGCAATTTACAGAGAGTGTAAAGTTAAACCCCAATTAAT 71779
QY 1380 CATGTTTATGATATGTAACATTAAGGCAATTTCAAAAAAGTGTCTAGGAGAGCT 1439
DB 71778 CATGTTTATGATATGTAACATTAAGGCAATTTCAAAAAAGTGTCTAGGAGAGCT 71719
QY 1440 CTTTGAAGAGTGGCAATTTGAGAGAGATGATGACAGAAAGCTAAACTGTGAG 1499
DB 71718 CTTTGAAGAGTGGCAATTTGAGAGAGATGATGACAGAAAGCTAAACTGTGAG 71659
QY 1500 TTTAAGGAGAAAGAAAGCACTGCAAAAGCCCTGAGCAGTAAAGATTTGCTGATT 1559
DB 71658 TTTAAGGAGAAAGAAAGCACTGCAAAAGCCCTGAGCAGTAAAGATTTGCTGATT 71599
QY 1560 CAAAGAAAGAGAGAAACCAATGCAACTGAGAAACAAAGTGGGCGCACATGTAAGAT 1619
DB 71598 CAAAGAAAGAGAGAAACCAATGCAACTGAGAAACAAAGTGGGCGCACATGTAAGAT 71539
QY 1620 GAGCTGAGAGTGTAGCAGAGGCGCAATGCTCTGCAAGTATTTCTTGTCTCAACACAG 1679
DB 71538 GAGCTGAGAGTGTAGCAGAGGCGCAATGCTCTGCAAGTATTTCTTGTCTCAACACAG 71479
QY 1680 AGCTTCCCTATGTTCTTAATGGAAGTGTATCTGTTGAGAAACAGAAATTTAAATCAAA 1739
DB 71478 AGCTTCCCTATGTTCTTAATGGAAGTGTATCTTGAAGAAACAGAAATTTAAATCAAA 71419
QY 1740 CTGTTATCATCAACAGACCTTCTCTGTAATCAAGCTCCCAAGGATCTTAAGAGAGCT 1799
DB 71418 CTGTTATCATCAACAGACCTTCTCTGTAATCAAGGCTCCCAAGGATCTTAAGAGAGCT 71359
QY 1800 AAGTTTAAACAAGCTCTCATTAAGCAGAGGTGTGTTTCAACAGTATTAAGAAAGCTGGGGAT 1859
DB 71358 AAGTTTAAACAAGCTCTCATTAAGCAGAGGTGTGTTTCAACAGTATTAAGAAAGCTGGGGAT 71299
QY 1860 TCAGAGTACTCCAGTCCCATGCTATGAAAAGCTCCCAATTTGTAACAACTGACA 1919
DB 71298 TCAGAGTACTCCAGTCCCATGCTATGAAAAGCTCCCAATTTGTAACAACTGACA 71239
QY 1920 AATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGCTCCGAGTGTGGGGGTGG 1979
DB 71238 AATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGCTCCGAGTGTGGGGGTGG -GG 71181

QY 1980 GTTCCGAGGGGAAACTTTTA 2001
 |||||
 Db 71180 GTTCCGAGGGGAAACTTTTA 71159

RESULT 9
 AC098800 142003 bp DNA linear PRI 15-OCT-2002
 LOCUS Homo sapiens BAC clone RP11-666p13 from 2, complete sequence.
 DEFINITION AC098800 AC068744
 AC098800.3 GI:23130726
 VERSION AC098800.3 GI:23130726
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 142003)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 142003)
 AUTHORS Shahid, S., Meyer, R., and Spalding, L.
 TITLE The sequence of Homo sapiens BAC clone RP11-666p13
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 142003)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 4 (bases 1 to 142003)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 5 (bases 1 to 142003)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (18-SEP-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 6 (bases 1 to 142003)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 7 (bases 1 to 142003)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (15-OCT-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 18, 2002 this sequence version replaced gi:22002231.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: saplens@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0666p13
 Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catalanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 This sequence is not the entire insert of the clone. This clone is overlapped by AC097532 and AC104405.

Uncertain bases from base 9817 to base 9869.

There is an unresolved tandem repeat from base 9870 to 11201.

Data from AC068744 and AC104405 was used to finish this clone.

The sequence of AC068744 has been incorporated into AC098800.

FEATURES

source

1..142003
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-666p13"
 /clone_lib="RPCT-11"
 1..154
 /rpt_family="Alu"
 410..518
 /rpt_family="MIR"
 1998..2193
 /rpt_family="L1"
 2251..2350
 /rpt_family="Achoho"
 2434..2605
 /rpt_family="MIR"
 3934..4288
 /rpt_family="L1"
 4289..4311
 /rpt_family="(TG)n"
 4312..5691
 /rpt_family="L1"
 5692..5717
 /rpt_family="(TTTTTA)n"
 5718..6180
 /rpt_family="L1"
 6804..7098
 /rpt_family="Alu"
 7139..7288
 /rpt_family="MER1_type"
 8352..8375
 /rpt_family="(GAAA)n"
 8397..8432
 /rpt_family="GA-rich"
 8569..8871
 /rpt_family="Alu"
 9165..9430
 /rpt_family="MALR"

```
repeat_region 9528..9727
/rpt_family="MER1_type"
repeat_region 12230..12519
/rpt_family="A1u"
repeat_region 13130..13267
/rpt_family="MER2_type"
repeat_region 13268..13806
/rpt_family="MER1_type"
repeat_region 13807..14000
/rpt_family="MER2_type"
repeat_region 14571..14710
/rpt_family="MIR"
repeat_region 14747..14772
/rpt_family="(TA)n"
repeat_region 14775..15038
/rpt_family="A1u"
repeat_region 15043..15760
/rpt_family="ERV1"
repeat_region 15838..15913
/rpt_family="u1"
repeat_region 16094..16263
/rpt_family="ERV1"
repeat_region 16266..16328
/rpt_family="MER1_type"
repeat_region 16339..16352
/rpt_family="(TAAA)n"
repeat_region 16370..16583
/rpt_family="MALR"
repeat_region 17083..17301
/rpt_family="MER1_type"
repeat_region 17470..17565
/rpt_family="u2"
repeat_region 17715..18240
/rpt_family="MALR"
repeat_region 18302..18341
/rpt_family="MALR"
repeat_region 18421..18816
/rpt_family="u1"
repeat_region 18980..19106
/rpt_family="A1u"
repeat_region 19141..19442
/rpt_family="A1u"
repeat_region 21116..21322
/rpt_family="(TTCC)n"
repeat_region 21328..23189
/rpt_family="u1"
repeat_region 23215..23315
/rpt_family="u6"
repeat_region 23316..23365
/rpt_family="u1"
repeat_region 23641..23897
/rpt_family="MIR"
repeat_region 23934..24124
/rpt_family="MIR"
repeat_region 25057..25137
/rpt_family="GA-rich"
repeat_region 25392..25488
/rpt_family="ISK"
repeat_region 25497..26487
/rpt_family="u1"
repeat_region 26514..26581
```

Query Match 14.3% Score 286.4; DB 9; Length 142003;
Best Local Similarity 72.5%; Pred. No. 3.3e-57;
Matches 429; Conservative 0; Mismatches 151; Indels 12; Gaps 4;

```
Qy 416 ATTTCCAGTTTATAGACCATTAACAGGTAACGCTGGGCGAGGCTCA 475
Db 27610 ACTTAATATGTAATAGACCATTAACAGGTAACGCTGGGCGAGGCTCA 27666
Qy 476 CACCTGTAATCCAGCACTTTGGAGGCTGAGGAGGAGATCACTTTGGTCAAGATT 535
Db 27667 CACCTGTAATCCAGCACTTTGGAGGCTGAGGAGGAGATCACTTTGGTCAAGATT 27726
```

```
Qy 536 TGAGACTAGCTGGCCCAACATGGCGAACTCTGCTCTAATAAATAATACAAATTAGC 595
Db 27727 CGAGACCACTGTAACCAATGCGAAACCC---CTCTCTACTAATAATATTAATAATTAT 27783
Qy 596 CTGGCATGTGTGGCGGCGCCTGTATCTCAGTGTCTCAGAGGCGGAGACAAAGATCA 655
Db 27784 CCGGCGTGTGTGGGCTGTATCTCTAATCTTAATCTTGGAGGCTGAGAGAGATCA 27843
Qy 656 CTTAAACCCAGAGGTGAGGTTGACGTGACGTGAGATCTGTCGACT-CACCTCAACTG 714
Db 27844 CTGGAACCCAGAGGTGAGGTTGACGTGAGGAGGAGATCTGTCGACTTTCAGGCTG 27903
Qy 715 GGAGACAGTGAACACTTTGTCTCAAAAAGAAAAAAGAAAAAGAAAAAGAAAG 774
Db 27904 GGTGACGTGATGAGACTCGCTCAAAAAAAGAAAAAAGAAAAAGAAAAATATTTT 27963
Qy 775 GTGAGTTAACTTAAATTAACCAATGATCCCAATCAATCAATTTCAAGTAAATTA 834
Db 27964 TAATTAATATTTTATTTATTAATCAATTAAGCAAAATATGATCATTTCAATGAAATCA 28023
Qy 835 TAATAAACAAATTAATGAATGAGATCTTACATTTCTTGTTCATATTAAGTCTTT 894
Db 28024 TATATAAACTAATTCAGAGATAATTACATTTCTT-----TTTTCATCAAAATCTTA 28078
Qy 895 GAAAGTGAATATATATTTATGCTGACAGACATCTCAATTTGAGACTAATTTTCAAG 954
Db 28079 GAAATCTGTCTCATTTTATTTATTTTCAAGGACATTTCAAGTTGAGACTAATTTTCAA 28138
Qy 955 GTGCTCAGTACCAATGTGGCTAGCAGTACTGATTTGATGAGCAGATC 1006
Db 28139 GTGTTTATTTAGTACAGCTGCTTAATGCTGATGATTAATAGCTCAGGCTC 28190
```

RESULT 10
AC133961/c 177787 bp DNA linear PRI 03-JAN-2003
LOCUS AC133961
DEFINITION Homo sapiens BAC clone RP13-494c23 from 4, complete sequence.
AC133961
VERSION AC133961.3 GI:26190601
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177787)
Sulston, J.E. and Watson, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 98063792
PubMed 9847074
REFERENCE
2 (bases 1 to 177787)
Iak, A., Bielecki, L., Creason, K. and Cotton, M.
TITLE The sequence of Homo sapiens BAC clone RP13-494c23
JOURNAL Unpublished (2001)
REFERENCE
3 (bases 1 to 177787)
Waterson, R.H.
AUTHORS Direct Submission
TITLE Submitted (20-SEP-2002) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
4 (bases 1 to 177787)
Waterson, R.H.
AUTHORS Direct Submission
TITLE Submitted (05-NOV-2002) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
5 (bases 1 to 177787)
Waterson, R.H.
AUTHORS Direct Submission
TITLE Submitted (07-DEC-2002) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE 6 (bases 1 to 177787)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2003) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Dec 7, 2002 this sequence version replaced gi:24580489.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@waterston.wustl.edu
----- Summary Statistics
Center project name: H_FH0494C23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCT-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01RG01165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.
Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1k2 or either MboI or PvuII for library segments 3k4. Size selected DNA was cloned into the pBACe3.6 vector between the EcoRI sites for library segments 1k2 or the BamHI sites for library segments 3k4. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

NEIGHBORING SEQUENCE INFORMATION:
The sequence from base 67288 to base 67340 was derived from one plasmid subclone.

Polymorphisms have been identified between AC020706 and this sequence.

Data from AC020706 was used to finish this clone.

This sequence is not the entire insert of the clone. This clone is overlapped by AC020706 and AC093660.

FEATURES

source

1. 177787
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP13-494C23"

repeat_region 4..70 /clone_11b="RPCT-13"
repeat_region /rpt_family="MIR"
repeat_region 703..833 /rpt_family="Alu"
repeat_region 852..875 /rpt_family="Alu"
repeat_region 911..1138 /rpt_family="AT-rich"
repeat_region /rpt_family="Alu"
repeat_region 1137..1191 /rpt_family="Alu"
repeat_region 1341..1408 /rpt_family="C(AAA)n"
repeat_region /rpt_family="GA-rich"
repeat_region 1432..1577 /rpt_family="GA-rich"
repeat_region 2115..2458 /rpt_family="GA-rich"
repeat_region /rpt_family="L2"
repeat_region 2562..2603 /rpt_family="L2"
repeat_region /rpt_family="MER1_type"
repeat_region 2773..2868 /rpt_family="MER1_type"
repeat_region /rpt_family="MER1_type"
repeat_region 2932..3240 /rpt_family="Alu"
repeat_region 3264..3462 /rpt_family="Alu"
repeat_region 3463..3760 /rpt_family="MALR"
repeat_region /rpt_family="Alu"
repeat_region 3761..3934 /rpt_family="MALR"
repeat_region /rpt_family="MALR"
repeat_region 5508..5796 /rpt_family="Alu"
repeat_region 6977..7240 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
repeat_region 7250..7270 /rpt_family="AT-rich"
repeat_region 7524..7661 /rpt_family="MIR"
repeat_region 7891..8423 /rpt_family="ERVL"
repeat_region 9624..9825 /rpt_family="MIR"
repeat_region 10105..10163 /rpt_family="L2"
repeat_region 10414..10721 /rpt_family="Alu"
repeat_region 11332..11671 /rpt_family="L1"
repeat_region 12367..12389 /rpt_family="AT-rich"
repeat_region 13174..13474 /rpt_family="Alu"
repeat_region 13507..13850 /rpt_family="Alu"
repeat_region 13834..13969 /rpt_family="Alu"
repeat_region 14004..14059 /rpt_family="polypyrimidine"
repeat_region 16122..16152 /rpt_family="T)n"
repeat_region 16824..17116 /rpt_family="Alu"
repeat_region 17122..17421 /rpt_family="Alu"
repeat_region 17422..18113 /rpt_family="Alu"
repeat_region /rpt_family="L1"
repeat_region 18681..18727 /rpt_family="L2"
repeat_region 20147..20162 /rpt_family="MIR"
repeat_region 20165..20386 /rpt_family="MER1_type"
repeat_region 20387..20502 /rpt_family="MIR"

| | Query Match | 13.8% | Score 276.8 | DB 9 | Length 17787 | |
|-----------------------|----------------|--|-------------------------|--------------|--------------|--|
| Repeat_region | 20503 | .20614 | /rpl_family="TA)n" | | | |
| repeat_region | 20617 | .20782 | /rpl_family="WIR" | | | |
| repeat_region | 21217 | .21523 | /rpl_family="L2" | | | |
| repeat_region | 21726 | .21824 | /rpl_family="CR1" | | | |
| repeat_region | 21947 | .22277 | /rpl_family="MER1_type" | | | |
| repeat_region | 24511 | .24634 | /rpl_family="L2" | | | |
| repeat_region | 24922 | .25153 | /rpl_family="MER1_type" | | | |
| repeat_region | 25154 | .25413 | /rpl_family="ALu" | | | |
| repeat_region | 25414 | .25462 | | | | |
| Query Match | 13.8% | Score 276.8 | DB 9 | Length 17787 | | |
| Best Local Similarity | 68.9% | Pred. No. 6.2e-55 | | | | |
| Matches 460 | Conservative 0 | Mismatches 187 | Indels 21 | Gaps 5 | | |
| Db | 311 | TCTGAGATCCAGCAATTTGCCAAGACAGTACAGACCCCTGTTCTCAAGAGCTCATACCC | 370 | | | |
| Db | 58503 | TCTGGCATATTTTCTTTTCCAGAAAGTACGCTGACCTCAATTTAGAAATGACATATC | 58444 | | | |
| Qy | 371 | TAGAGTAGTGTTTGTAGTAAATTAATGCTGAGCTGCTTATGTCATTTCCAGTTTCTTA | 430 | | | |
| Db | 58443 | CAGTTGAAATACATATGTAACCAAAATTTCAAGCCACATATATCATTTTAAATTCCTTA | 58384 | | | |
| Qy | 431 | GTAACCATTTAAACAGCTAAAAAAGGCTGGGCGAGTGGCTCACCTGTAAATCCGAG | 490 | | | |
| Db | 58383 | GTACTCTATGTTAAAAAGTAAAGAGGCGCGGGGAGTGGCTCACCTGTAAATCTTGA | 58324 | | | |
| Qy | 491 | CACTTTGGGAGGCTGAGGACAGGACATCACCTTTGTCAAGAGTTTGAAGTACGCTGAG | 550 | | | |
| Db | 58323 | CACTTTGGGAGGCGGAGGAGGAGTGGATCACCTGAGTCAAGAGTTTGAAGACGAGCTGAG | 58264 | | | |
| Qy | 551 | CAACATGCGCAACTCTGTCTCTAAAAAAATTAAGCTGAGTGGTGGCGG | 610 | | | |
| Db | 58263 | CAACATGCGCAAAACCTGTCTCT-ACGTAAATTAAGCTGGGAGTGTGGCAC | 58205 | | | |
| Qy | 611 | GGCGCTGTAAATCTCAAGCTGCTCAGAGGCGGAGACAAAGATCATTAATCCAGAGG | 670 | | | |
| Db | 58204 | ATGCGCTGTGATCTAGCTGCTCAGAGGCTGAGGACAGAAATCGCTAGAAACCGGGAGG | 58145 | | | |
| Qy | 671 | TGAGAGTTGACGAGGAGTGAATGTCGCACT-CATCCAACTGGGAGCAGAGTGACA | 729 | | | |
| Db | 58144 | CGGAGTTGCAATGAGCGGAGATTTGCCCTTTGTGACTCCAGCTGGGCGCAAGAGGACA | 58085 | | | |
| Qy | 730 | CTTTTGTCTCAAAAGAAAAAAGAAAAAAGAAAAAGAAAGAGTGAATTAATCTTTA | 789 | | | |
| Db | 58084 | CTTCATCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAGTAAAGAGCGCATTAATGAAT | 58025 | | | |
| Qy | 790 | ATAACCAATGAT-----CCCAATACATCATTTCAAAGTGAATTAATA | 836 | | | |
| Db | 58024 | TTAATGTATTTAATTAACACAGTATCCAAATATGTATGATCAACATCAATTAATA | 57965 | | | |
| Qy | 837 | TAAACAATTTGAATGAGTACTTCAATTTCTTTCTTTTCAATTAATTAAGCTTTGA | 896 | | | |
| Db | 57964 | TAAA-ATTATTAATGATGATATTTTCCAGCC-----CTTTTTTTAACAAGGTTTGA | 57911 | | | |
| Qy | 897 | AAGTAGATATATGTTATGCTGACAGACATCTCAATTTGAGTACGCTACATTTCAAGT | 956 | | | |
| Db | 57910 | AATTCATGATATTTTACCTTACAGAGATCTCAATGAGACGAGCCACATTAACAAT | 57851 | | | |
| Qy | 957 | GCTCAGTA 964 | | | | |
| Db | 57850 | GCTTAATA 57843 | | | | |

| LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT | | | | | | | | | | | | | | | | | |
|------------|---|---------------------|-----------|--------------------------|-------------------------|-----------------------|-----------------------|--------------------|---------------------|---|--|------------------------|---------------------------------|------------------------|---------------------------------|------------------------|---------------------------------|------------------------|-------------------------------|------------------------|--------------------------------|------------|---------------------------------|------------|---------------------------------|------------|--|------------|
| HS007J7H19 | Human DNA sequence from clone RPI-7JH19 on chromosome 6 Containing part of the PRIM2A gene encoding DNA primase large subunit (p55), ESTs, STS and GSSs, complete sequence. | U0121958 | 6 | GI:8247268 | HTG; PRIM2A. | Homo sapiens (human) | 1 (bases 1 to 138411) | Brown, J. | Direct Submission | Submitted (27-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hunknery@sanger.ac.uk | <p>Requests: clonerequests@sanger.ac.uk</p> <p>On Jun 4, 2000 this sequence version replaced gi:8218079.</p> <p>During sequence assembly data is compared for overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.</p> <p>This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unseq' feature key.</p> <p>The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpeg This sequence was generated from part of bacterial clone conligs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6</p> <p>RPI-7JH19 is from the library RPCI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/VECTOR:PCYRAC2</p> <p>This sequence is the entire insert of clone RPI-7JH19 The true left end of clone RPI-216J23 is at 87497 in this sequence. The true right end of clone RPI-422B11 is at 29670 in this sequence.</p> | | | | | | | | | | | | | | | | | |
| FEATURES | SOURCE | Location/Qualifiers | 1..138411 | /organism="Homo sapiens" | /mol_type="genomic DNA" | /db_xref="taxon:9606" | /chromosome="6" | /clone="RPI-7JH19" | /clone_lib="RPCI-1" | complement(477..1095) | /note="match: GSS: Em:AQ0506907" | complement(609..11046) | /note="match: GSS: Em:AQ437147" | complement(620..11089) | /note="match: GSS: Em:AQ819330" | complement(621..11083) | /note="match: GSS: Em:AQ437147" | complement(761..11000) | /note="match: GSS: Em:BM1825" | complement(977..10930) | /note="match: GSS: Em:AQ55044" | 1514..1858 | /note="match: GSS: Em:AQ176950" | 1545..1851 | /note="match: GSS: Em:AQ176950" | 1857..2161 | /note="AluX repeat: matches 1..309 of consensus" | 2510..8661 |

```
misc_feature /note="L1PA4 repeat: matches 13. .6142 of consensus"
8448. .8976
/note="match: GSS: Em:AQ347835"
misc_feature 8662. .8847
/note="match: GSS: Em:AQ347835"
8648. .9156
repeat_region /note="MLT11 repeat: matches 40. .410 of consensus"
9172. .9481
/note="AluSq repeat: matches 5. .313 of consensus"
9692. .9804
/note="MLT11 repeat: matches 267. .371 of consensus"
9824. .10043
/note="MIR repeat: matches 18. .251 of consensus"
10338. .10689
/note="MLT2FA repeat: matches 9. .356 of consensus"
complement(10977. .11390)
misc_feature /note="match: SRS: Em:G42414"
complement(11825. .12306)
/note="match: GSS: Em:AQ141424"
12402. .12825
/note="match: GSS: Em:AQ266412"
12647. .12722
repeat_region /note="MIR repeat: matches 1. .79 of consensus"
13978. .14287
/note="AluDb repeat: matches 1. .310 of consensus"
14664. .15035
/note="L2 repeat: matches 2294. .2673 of consensus"
15036. .15342
/note="AluSp repeat: matches 1. .308 of consensus"
15343. .15421
/note="L2 repeat: matches 2673. .2750 of consensus"
15485. .15534
/note="MIR repeat: matches 213. .262 of consensus"
15581. .15703
/note="MIR repeat: matches 110. .234 of consensus"
16310. .16535
/note="MIR repeat: matches 21. .231 of consensus"
16597. .17598
/note="L1PA7 repeat: matches 5135. .6145 of consensus"
17617. .17827
/note="AluSp repeat: matches 83. .309 of consensus"
17828. .19349
/note="L1PA7 repeat: matches 3624. .5146 of consensus"
complement(19723. .20194)
misc_feature /note="match: GSS: Em:AQ016849"
20519. .20665
/note="match: GSS: Em:AQ371979"
20807. .21136
/note="match: GSS: Em:AQ508843"
20807. .21096
/note="match: GSS: Em:AQ508843"
20818. .21272
/note="match: GSS: Em:AQ209413"
20818. .21096
/note="match: GSS: Em:AQ209413"
20831. .21139
/note="match: GSS: Em:AQ514718"
20902. .21139
/note="match: GSS: Em:AQ093115"
21097. .21176
/note="40 copies 2 mer ct 76% conserved"
21524. .22065
/note="match: GSS: Em:AQ50328"
22856. .23094
repeat_region /note="AluSq repeat: matches 1. .239 of consensus"
23102. .23135
/note="17 copies 2 mer aa 88% conserved"
24117. .24204
/note="MER5B repeat: matches 1. .90 of consensus"
24883. .25185
/note="AluY repeat: matches 1. .300 of consensus"
25364. .25449
/note="MSTD repeat: matches 1. .96 of consensus"
```

```
repeat_region 25932. .26300
/note="L1MB8 repeat: matches 5779. .6162 of consensus"
repeat_region 26483. .26600
/note="L2 repeat: matches 2583. .2703 of consensus"
27463. .27633
/note="AluSp repeat: matches 137. .308 of consensus"
misc_feature complement(27586. .28035)
/note="match: GSS: Em:AQ753792"
complement(27650. .27736)
/note="match: GSS: Em:AQ753792"
27737. .28017
/note="AluDb repeat: matches 1. .276 of consensus"
28766. .28860
/note="MIR repeat: matches 85. .188 of consensus"
29043. .29225
/note="PRAM repeat: matches 0. .174 of consensus"
29387. .29497
/note="MIR repeat: matches 29. .143 of consensus"
29585. .29879
/note="AluSp repeat: matches 1. .296 of consensus"
31001. .31343
/note="L1MD2 repeat: matches 5957. .6332 of consensus"
31344. .31643
/note="AluSc repeat: matches 1. .301 of consensus"
31644. .31712
/note="L1MD2 repeat: matches 5890. .5957 of consensus"
31714. .31863
/note="L1MD3 repeat: matches 7589. .7735 of consensus"
31874. .32598
/note="L1MD2 repeat: matches 5386. .6122 of consensus"
33641. .33934
/note="AluXk repeat: matches 3. .298 of consensus"
33933. .33941
/note="single clone region"
34941. .35324
/note="match: GSS: Em:B55255"
35189. .35758
/note="L2 repeat: matches 2189. .2747 of consensus"
complement(36539. .37046)
/note="match: GSS: Em:AQ558962"
36809. .37468
/note="match: GSS: Em:AQ541391"
36811. .37282
/note="match: GSS: Em:AQ421653"
36831. .37467
/note="match: GSS: Em:AQ547524"
36992. .37051
/note="MER85 repeat: matches 86. .139 of consensus"
37054. .37177
/note="MER85 repeat: matches 1. .139 of consensus"
37956. .38473
/note="match: GSS: Em:AQ246505"
complement(37968. .38506)
/note="match: GSS: Em:AQ427574"
misc_feature
```

Query Match 13.7%! Score 274.4; DB 9; Length 138411;
Best local Similarity 78.0%; Pred. No. 2.3e-54;
Matches 415; Conservative 0; Mismatches 91; Indels 26; Gaps 6;

```
QY 472 CTCACACCTGTATATCCAGCACTTTGGAGCTGAGCGAGCAATCACCTTGGTCAGG 531
|||||
DB 43069 CTCACACCTGTATATCTCAGCACTTTGGAGCGCGAGCGAGATCACTGAGGTCAGG 43128
|||||
QY 532 AATTGAGACTAGCTTGCGCAACATGGGAAATCTGTCTCTCAAAAAAATACAAAAT 591
|||||
DB 43139 AATTGAGACTAGCTTGCGCAACATGGGAAATCTGTCTCT-CTTAATAAATACAAAAT 43187
|||||
QY 592 TAGCTGCGATGTGGCGGCGGCTGTATATCTCAGCTGTCAGAGCGGAGACAAGA 651
|||||
DB 43188 TAGCGGGCATGTGTGGGTAATCTGTATATCCAGCTACTCAGAGGCTGAGGCGAGGA 43247
|||||
QY 652 ATCACTTAACCCAGAGGTGAGGTTGCACTGAGCTGATGTCGCACT-CACTCCAA 710
|||||
```

| | | | |
|----|-------|--|-------|
| Db | 43248 | ATTGCTTTGAACCCAGGAGGCGAGGGCTGACGTAGGCCAGATTGTGCACCTGCAC | 43307 |
| Qy | 711 | CTTGGGAGCAGAGTGCACCTTTGTCTCAAAAAAAAAAAAAAAAACAAGTAAAAAGAA | 770 |
| Db | 43308 | CTTGGGTGACAGAGGAGAACTCTCTTAAAAAAAAAAAAAAAAAAAAAGAAAAA | 43367 |
| Qy | 771 | ACAGG-----TGAAGTTAACTTTAATPACCAGATGATCCAAATACATCATTTTC | 821 |
| Db | 43368 | GGATGAAGTAAATTTTATTTATTTGATTTTAAACCAATATATTGCAAATATGTGATTTT | 43427 |
| Qy | 822 | AAATGTATATTAATATTAACAATATATGATGATPACTTTACATTTCTTTGTGTTTC | 881 |
| Db | 43428 | GATATATATATATAGTATCAAAAGTT---ATTGGAATATTTTAACTCTTTTCTCT--- | 43478 |
| Qy | 882 | ATATTAAAGTCTTTGAAAGTAGATATATATGTATATGTGCACAGACATCTCAATTTGACT | 941 |
| Db | 43479 | ATGCTAAGTCTTTGAAATTTTGTGTATATTTTATG-----GTATATCTCAATTTGACT | 43522 |
| Qy | 942 | AGCTACATTTCAAGGTGCTCAGTAGGCCACATGTGGCTACAGAGTAACTGTATGG | 993 |
| Db | 43533 | AGCCACATTTCAAGATCTCAGTAGCCACATGTGGCTATAGTACGTATATG | 43584 |

| | |
|------------|--|
| RESULT | 12 |
| LOCUS | AC087591/c |
| DEFINITION | Homo sapiens chromosome 3 clone RP11-95M5 map 3p, complete sequence. |
| ACCESSION | AC087591 |
| VERSION | AC087591.2 |
| KEYWORDS | HTC. |
| SOURCE | Homo sapiens |
| ORGANISM | Homo sapiens (human) |

| REFERENCE | AUTHORS |
|--------------------------------|--|
| 1 (bases 1 to 161014) | Zhang, H., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Deng, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, N., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, L., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H. |
| Chromosome 3p genomic sequence | Unpublished |

TITLE Direct Submission
JOURNAL Submitted (12-JAN-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, 100101, P. R. China
REFERENCE 3 (bases 1 to 161014)
AUTHORS Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, Y., Ding, H.,

3 (bases 1 to 161014)

REFERENCE
AUTHORS

Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y.,

TITLE
JOURNAL
 Mr.D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
 Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
 Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
 Yu,J. and Yang,H.
Direct Submission
 Submitted (08-NOV-2002) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China
 4 (bases 1 to 161014)
REFERENCE
AUTHORS
 Zhang,H., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
 Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q.,
 Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N.,
 Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y.,
 Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q.,
 Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R.,
 Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X.,
 Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C.,
 Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,L., Zhang,M.,
 Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
 Yu,J. and Yang,H.
Direct Submission
 Submitted (20-DEC-2002) Human Genomic Center, Institute of

100101, P.R.China
On Nov 8, 2002 this sequence version replaced gi.12084032.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
-----Project Information
Center project name:1# project

```

FEATURES
SOURCE
-----
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 4% of reads
Assembly program: Phrap; version 0.990129
Consensus quality: 162662 bases at least Q40
Consensus quality: 163749 bases at least Q30
Consensus quality: 164140 bases at least Q20
Insert size: 161014; sum-of-contigs
Quality coverage: 9.63x in Q20 bases,sum-of-contigs
-----
Location/Qualifiers
1. 161014
/organism="Homo sapiens"
/mol_type="genomic DNA"

```

```

Db      98725  ACATCTAACCTTACTATCTCTTAATCCAGGCGTGTCCCTGGAAATGTAGGCACATATATC 9886
Qy      404    GCTGCTATTATCATTTCCAGCTTTTATAGACCAATT-AAAAAGGTAAAAAGCGCTGG 462
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      98865  ACTACATAGGCAAATTTAAATTATCTTGTAGGCACATTAAAAATTAATATAGCTAAGCGCTGG 98960

```

Db 98865 AACTACATAGGCATTTTAAATTCTTGTAAGCACAATAAAATAATAGTAAAGCTGG 98860

Qy 463 GCGCAGTGSGCTCAACCTGTAATCCAGCACTTTGGAGGCTGAGCGCAGCAGATCACT 522

Db 98865 GTGCGAGTGCGCTCACACATGTAATCCAGCACTTTCAGAAAGCTAAAGGGCAGATCACT 98854

Qy 523 TTGGTCAGGAGTTTGAGACTAGCTGGCCAACTGGCGAAACTCTGTCTTAATAAAAAA 582

Db 98545 GAGGTGAGAGTTGAGACGACCGCTGTCAACATGCTGTAACCCCGTCTCT-ACATAAAA 98487
Qy 583 TACAAATAATGAGCTGAGATGAGTGGGCGGCGCTGTATCTGAGTCTGAGAGGCGCA 642
Db 98486 CACAAATAATGAGCTGAGATGAGTGGGCGGCGCTGTATCTGAGTCTGAGAGGCGCA 98427
Qy 643 GACACAGAATCACTTAAACCCAGAGGTGGAGTGGTCAAGTACAGTCCGTGCACT 702
Db 98426 GGCAGAGGATCACTTAAACCTGGAGACAGAGGTTGCATATAGCCAAAGTTAGGCACT 98367
Qy 703 -CACTCAACCTGGAGACAGAGTGCACCTTTGCTCAAAAAAGAAAAACAAGT 761
Db 98366 GAACCTCCGCTGGGAGACAGCAAACTCCGTCTCAAAAAAGTAAAGTATTTT 98307
Qy 762 AAAAAAGAAACAGGTGAGTAACTTAAATACCAATGATCCCAATATCAATCATTTTC 821
Db 98306 AACGA-----TATATTTTATTTGAGCTCAATATATCAAAATTTATTCATTTTC 98261
Qy 822 AAAGTGAATTAATTAATTAACCAATTAATGAGATGAGTAACTTTCTTTTCTTTTC 881
Db 98260 AATATGTAATCAATATATTAATAATGTAATGAGTGTGACACTT-----TTTTC 98207
Qy 882 AATATTAAGTCTTGAAGAGATATATGTTATGCTGACAGACATCTCAATTTGCACT 941
Db 98206 AATCTAAGTGTCCCAATCTTAATATGCAATTTTACATTTGACAGCACTTCAATTTAGACT 98147
Qy 942 AGCTACATTTGAGTGTGCTGCTGAGTCCAGATGCTGAGTAACTGTAATTTGAGTGGAC 1001
Db 98146 AGCTACATTTTACCTTCTCTCAACAGCAATGCGGCTGTGGCTAGCGTACTGTCCAGCAC 98087
Qy 1002 GGAATCTAGAGGG 1013
Db 98086 AGCTCTAAGTGG 98075

RESULT 13
AC025163
LOCUS
DEFINITION
AC025163 177876 bp DNA linear HTG 26-JUN-2001
Homo sapiens chromosome 3 clone RP11-533F2, WORKING DRAFT SEQUENCE,
17 unordered pieces.
AC025163
VERSION
AC025163.17 GI:14547360
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177876)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benson,J., Binsag,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dabhorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Dem,A.L., Ding,Y., Dinb,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Recha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garrar,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korval,J.,
Kovar,C., Kravovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,U., Li,Z., Lichtarge,O., Lieu,C., Liu,X., Liu,W.,
Lousaged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,B., Massey,B., Mawliny,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenwo,S.,
Oguth,M., Okunoyu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Prims,E., Pu,L.L.,
Quiles,M., Ren,Y., Rivers,M., Roja,A., Rojudoan,I., Rolle,M.,
Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshari,N.,
Stison,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Stanley,H.,
Tang,H., Tansy,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watling,K., Wu,C., Williams,G., Williamson,A., Wiczynk,R., Woodson,S.,
Worley,K., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 177876)
Worley,K.C.
Direct Submission
Submitted (07-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:12831249.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAMH
Center clone name: RP11-533F2
Sequencing vector: M13; L08821
----- Summary Statistics
Chemistry: Dye-terminator Big Dye 3.1
Assembly program: Phrap; version 0.990329
Consensus quality: 169802 bases at least Q40
Consensus quality: 180431 bases at least Q30
Consensus quality: 185136 bases at least Q20
Estimated insert size: 177736; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-1p estimation
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
42878: contig of 42878 bp in length
42879
42978: gap of unknown length
42979
62018: contig of 19040 bp in length
62019
62118: gap of unknown length
62119
80603: contig of 18485 bp in length
80604
80703: gap of unknown length
80704
98288: contig of 17585 bp in length
98289
98388: gap of unknown length
98389
113934: contig of 15546 bp in length
113935
114034: gap of unknown length
114035
125339: contig of 11305 bp in length
125340
125439: gap of unknown length
125440
138329: contig of 12890 bp in length
138330
138429: gap of unknown length
138430
145677: contig of 7148 bp in length
145678
145679: gap of unknown length
145680
150952: contig of 5275 bp in length
150953
151052: gap of unknown length
151053
155923: contig of 4871 bp in length
155924
156023: gap of unknown length

* 159024 159748: contig of 3725 bp in length
 * 159749 159848: gap of unknown length
 * 159849 163146: contig of 3298 bp in length
 * 163147 163246: gap of unknown length
 * 163247 166618: contig of 3372 bp in length
 * 166619 166718: gap of unknown length
 * 166719 170864: contig of 4146 bp in length
 * 170865 170964: gap of unknown length
 * 170965 173099: contig of 2135 bp in length
 * 173100 173199: gap of unknown length
 * 173200 175517: contig of 2318 bp in length
 * 175518 175617: gap of unknown length
 * 175618 177876: contig of 2259 bp in length.

FEATURES
 Source
 1. 177876
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-533F2"

ORIGIN

Query Match 13.2%; Score 264; DB 2; Length 177876;
 Best Local Similarity 67.6%; Pred. No. 6.8e-52;
 Matches 454; Conservative 0; Mismatches 195; Indels 23; Gaps 5;

QY 344 ACCCTGTTCTACAGAGCTCAATACCTAGAGTAGGTGTTAGTGAATAATGCTGA 403
 DB 22489 ACATCTAACATTAATCTAATCTAATCAAGGCTGTCCCTGGATGTAGCCACATGATC 22548
 QY 404 GCCTGTAATGCTATTCAGTTTGTAGTGGCCACTT-AAAGAGTAAAGAGTGG 462
 DB 22549 ACTACATAGCAATTTTAAATATCTTGAGGCACATTAATAATTAATAGTAAAGCTGG 22608
 QY 463 GCCAGTGGCTCACACCTGTATCCAGCACTTTGGAGGCTGAGGAGGAGATCACT 522
 DB 22609 GTGCAAGTGGCTCACACATGTATCCAGCACTTCAAGAGTAAAGTGGGAGATCACT 22668
 QY 523 TTGGTCAGAGTTTGAAGACTAGCTGGCCAAATGCGGAACTGTGCTCTTAAAAAAA 582
 DB 22669 GAGGTGAGAGTTGAGAGCCAGCTGGTCAACATGTGAAACCCCGTCTCT-ACTAAAAA 22727
 QY 583 TACGAAAAATTAAGCTGCAATGCTGCGGCGGCGCTGTATCTCACTGCTCAGAGGCGCA 642
 DB 22728 CAGAAAAATTAAGCGGTGTGTGTGACACGCTGTAGTCTCACTACTCAGAGGCTGA 22787
 QY 643 GACACAGAAATCACTTAAACCCAGAGGTGAGGTTGAGTGAAGTCCGCACT 702
 DB 22788 GGCAGAGAAATCACTTAAACCTGGAGACAGAGTTGCATTAAGCCAAAGATGAGCACT 22847
 QY 703 -CACTCCAACTCGGAGACAGAGTGAACCTTTGTCTCAAAAAAGAAAAAACAAGT 761
 DB 22848 GAACCTCCGCGCTGGGTGACAAAGCAAAACCTCGTCAAAAAAGTAAATTTT 22907
 QY 762 AAAAAAAGAAAGGTGAAGTTAACTTTAATACCAGATATCCCAATATACATCACTTC 821
 DB 22908 AACGA-----TATATTTTATGGAATCAATATATCAATATATTCATTTTC 22953
 QY 822 AAAGTGAATTAATATATAAACAATTAATGAATGAGATCTTACATCTTTCTGTTTC 881
 DB 22954 AATATGTAATCAATATATAAATATTAATGAAGTGTGACACTCTT-----TTTTC 23007
 QY 882 ATATTAAGTCTTTGAAAGTGAATATATGTTATGCTGACAGACATCTCAATTTGACT 941
 DB 23008 ATACTAAGTGTCCCAATCTAATATGCAATTTTATGACAGACATCTCAATTTGACT 23067
 QY 942 AGTACATTTTCAAGTGTCTGAGTACCAATGTGCTGCACTTCTGATTTGATGAGAC 1001
 DB 23068 AGTACATTTTATCTGCTCAAGAGCCACATGCGCTGTGCTAGGTAATGTCACGAC 23127
 QY 1002 GGATCTAGAGGG 1013
 DB 23128 AGCTTAAGTGG 23139

COMMENT

RESULT 14
 ALJ35815
 LOCUS
 DEFINITION
 ALJ35815 92882 bp DNA linear PRI 21-DEC-2000
 Human DNA sequence from clone RP3-453H5 on chromosome 6. Contains
 ESTs, STSs, GSSs and a Cpg island. Contains a novel gene, complete
 sequence.
 ACCESSION
 ALJ35815
 VERSION
 ALJ35815.9 GI:10443432
 KEYWORDS
 HTG; Cpg island.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 92882)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Parker, A.
 Direct Submission
 Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Oct 1, 2000 this sequence version replaced gi:10186513.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chrg
 IMPORTANT: This sequence is not the entire insert of clone
 RP3-453H5. It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP3-453H5 is at 1 in this sequence. The
 true left end of clone RP3-470K1 is at 92783 in this sequence. This
 sequence has been finished according to sequence map criteria as
 follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated repeat sequence elements. Where the sequence is
 ambiguous, there is an annotation using the 'unsure' feature key.
 RP3-453H5 is from the library RPCI-3 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2.

FEATURES

Source
 1. 92882
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP3-453H5"
 /clone_1lb="RPCI-3"
 4. 33
 /note="Alu repeat: matches 26.55 of consensus"
 29.566
 /note="match: STS: Em:HS453H5"
 2269.2693
 /note="LIME2 repeat: matches 5523.5962 of consensus"
 2764.2993
 /note="LIME2 repeat: matches 5141.5391 of consensus"
 2994.3398
 /note="LIMD3 repeat: matches 7338.7740 of consensus"
 3399.3414
 /note="LIME2 repeat: matches 5128.5141 of consensus"
 3673.4100
 /note="L1 repeat: matches 4439.4840 of consensus"

```
repeat_region 4101. .4399
/note="A1uSx repeat: matches 1. .299 of consensus"
repeat_region 4400. .4765
/note="L1 repeat: matches 4004. .4439 of consensus"
repeat_region 4772. .5034
/note="L1M4 repeat: matches 3188. .3469 of consensus"
repeat_region 5035. .5355
/note="A1uO repeat: matches 1. .309 of consensus"
misc_feature complement(5142. .5568)
/note="match: GSS: Em:B71585"
misc_feature complement(5243. .5586)
/note="match: GSS: Em:B71584"
repeat_region 5356. .5395
/note="L1M4 repeat: matches 3151. .3188 of consensus"
repeat_region 5942. .6441
/note="L2 repeat: matches 1484. .2055 of consensus"
repeat_region 7093. .7168
/note="MIR repeat: matches 97. .174 of consensus"
repeat_region 7381. .7748
/note="MER47A repeat: matches 1. .365 of consensus"
repeat_region 8446. .8516
/note="MIR repeat: matches 68. .141 of consensus"
repeat_region 9113. .9216
/note="26 copies 4 mer agag 66% conserved"
repeat_region 9146. .9217
/note="36 copies 2 mer ga 77% conserved"
repeat_region 9428. .9789
/note="L1M4C repeat: matches 1169. .1539 of consensus"
repeat_region 11328. .11428
/note="MER5A repeat: matches 9. .112 of consensus"
repeat_region 11564. .11680
/note="L2 repeat: matches 2581. .2710 of consensus"
repeat_region 12030. .12247
/note="MER21B repeat: matches 414. .627 of consensus"
repeat_region 12569. .12622
/note="MUT2B repeat: matches 394. .448 of consensus"
repeat_region 12623. .12658
/note="18 copies 2 mer ta 100% conserved"
repeat_region 12727. .13139
/note="MUT2B repeat: matches 1. .404 of consensus"
repeat_region 13140. .18697
/note="HERVL repeat: matches 1. .5654 of consensus"
repeat_region 18698. .19189
/note="MUT2B repeat: matches 1. .444 of consensus"
repeat_region 19191. .19821
/note="MER21B repeat: matches 146. .785 of consensus"
repeat_region 20231. .20491
/note="MUT1A1 repeat: matches 105. .365 of consensus"
repeat_region 21266. .21603
/note="MUT1A1 repeat: matches 3. .364 of consensus"
repeat_region 21917. .22261
/note="MER1B repeat: matches 1. .337 of consensus"
gene complement(22369. .44784)
/note="product="d4453H5.1" (novel protein)"
mRNA match: ESTs: Em:AW02079 Em:AA509875 Em:AV728994 Em:N91539
Em:BE906111 Em:BE372189 Em:AV725475 Em:AA229477
Em:BF031801 Em:AV659500 Em:AA357194 Em:AA075033
Em:AA228721 Em:W38714 Em:AA357193 Em:CO3105 Em:AA091429
Em:AA936129 Em:AW605144 Em:AA605142 Em:AW015656
Em:AA72673 Em:AI680140 Em:AA832355 Em:AA491183
Em:BE645240 Em:AA169452 Em:AA971488 Em:AW197692
Em:AA618518 Em:BE815910 Em:AA085746 Em:AI620439
Em:AI420298 Em:AI936423 Em:AA248996 Em:AA658139
Em:BE815919 Em:BE697434 Em:BE676861 Em:BE697435
Em:BE697436 Em:AW655902 Em:AM130169 Em:AI473966
Em:AW615082 Em:AW276216 Em:AW197858 Em:AW786005
Em:AA767728 Em:AA075034 Em:AA935887 Em:AA831190
```

```
Em:AI769381 Em:AI302136 Em:AW975096 Em:BE247082
Em:AI138958 Em:AA395621 Em:AW301242 Em:AW276204
Em:BE697440 Em:AA827775 Em:BE697442 Em:AI910858 Em:H04554
Em:AA471313 Em:AI765568 Em:AV729224 Em:AW633777 Em:W07150
Em:AA652231 Em:AW054987 Em:AI955586 Em:AA401305 Em:N92598
Em:AI445162 Em:BE048962 Em:AI245709 Em:AI915354
Em:AI955631 Em:BE754575 Em:AA79816 Em:AW630465
Em:AW197776 Em:AW467500"
/evidence=not experimental
polyA_site complement(22369)
/gene="d4453H5.1"
polyA_site complement(22371)
/gene="d4453H5.1"
polyA_site complement(22372)
/gene="d4453H5.1"
polyA_site complement(22373)
/gene="d4453H5.1"
polyA_site complement(22375)
/gene="d4453H5.1"
polyA_signal complement(22392. .22397)
/gene="d4453H5.1"
CDS complement(join(22850. .23000,24209. .24278,26185. .26296,
26703. .26795,28881. .28978,30022. .30129,31771. .31925,
37238. .37359,39165. .39458,43686. .43688))
/gene="d4453H5.1"
/codon_start=1
/evidence=not experimental
/product="d4453H5.1 (novel protein)"
/protein_id="CAC19504.1"
/db_xref="GI:11990034"
/db_xref="UniProt/TREMBL/Q9H1N7"
/translation="MDLTQDAKIDQNIITVOETNNKNSBSIBCSKITMDLKPNNRSKYI
SITVPSKQSGWPHISKVDVVDVILMNISKEFKLQFICVAGVFPVLIYGLDELI
PSVREKSCQWTLTLPORAPYSIPGLTLOLIDRRRIPGTWMTIAFLYGTMGSL
NTSLGLYLPYQVITPKCKLIPWMLGCFITQSKRNVDVSAITMSGLIYPTLADS
TTAFPNLTGVVLISLALCAADVIGNVBERAKKLNASNSEKVLTSIGFYIILGL
TCTSGIAPAVTFECANPRTYGAFLFSLTGYFGLSVIALIKIFGALIAVTTGK
AMTIVLSIFEFKAPTFQYVWSGLVLVIGIPLMYVSKNMDKIRLPSLYDLINKSEAR
KSRTLAQTV"
repeat_region 24052. .24185
/note="67 copies 2 mer aa 58% conserved"
repeat_region 24785. .24986
/note="L2 repeat: matches 2477. .2710 of consensus"
repeat_region 25084. .25159
/note="MUT1A1 repeat: matches 107. .186 of consensus"
repeat_region 25227. .25288
/note="MUT1A1 repeat: matches 326. .396 of consensus"
repeat_region 25296. .25464
/note="MIR repeat: matches 2. .188 of consensus"
repeat_region 27094. .27187
/note="47 copies 2 mer aa 63% conserved"
repeat_region 27313. .27538
/note="MIR repeat: matches 3. .235 of consensus"
repeat_region 27770. .27801
/note="16 copies 2 mer ac 93% conserved"
repeat_region 27772. .27803
/note="8 copies 4 mer acac 93% conserved"
repeat_region 28055. .28139
Query Match 13.1%; Score 262.4; DB 9; Length 92882;
Best Local Similarity 71.8%; Pred. No. 1,6e-51;
Matches 428; Conservative 0; Mismatches 156; Indels 12; Gaps 6;
QY 370 CTAAGTAGTGCTGTTTACTAGAAATTA-ATGCTGAGCTCCTTATGCTATTCAGATTGTTT 428
DB 50297 CCAGAACAGTGTCTGCATATAGAAATTAATTAATTAAGCCATGGAACCTACTTAAATTTTC 50356
QY 429 TAGTAGCCACATTAAACAGTAATAAAGGCTGGGCGCAGTGGCTCACACTGTATCCC 488
DB 50357 TCTTAGTACATTAAAGAGTAATAATGAGCGCGGCTGTGCTGCTACATGCTCTATGCC 50416
QY 489 ACCACTTTGGAGGCTGAGCAGGAGATCACCTTTGTCAGAGCTTGAGACTAGCCTTG 548
|||||
```

Db 50417 AGCACTTTGGAGGCGCAAGGCGAATCATCTTGGAGTCAGAGTTCAGAACCCAGCTG 50476
 QY 549 GCCAATCTGGCGAAACCTCTGCTCTA-AAAAAAAAATACAAAATTAAGCTGGATGGTG 607
 Db 50477 GCCAATCTGGCGAAACCCCTCTCTACTAATAAATAAATAAATTAACAGGCGCTGTGG 50536
 QY 608 CGGGCGGCTGTAACTCTAGCTGCTCAGAGGCGCAGACAAAGATCACTTAAACCAAG 667
 Db 50537 CGGGCGGCTGTAACTCTAGCTGCTCAGAGGCGCAGAGATGAGTCTTAAACCAAG 50596
 QY 668 AGGTGAGGTTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 726
 Db 50597 AGGCAAGGTTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 50656
 QY 727 ACACCTTTCTCTCAAAAAGAAAAAAGAAAGTAAAGAAAGAAAGTGAAGTGAAGT 786
 Db 50657 AGTCTCTGTCTCAAAAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 50712
 QY 787 TTAATAACCAATGTATCCCAATATACATCTTCAAGTGAATTAATTAATTAATTAATTAAT 846
 Db 50713 TATTAACTCAATTAACCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 50771
 QY 847 ATGATGAGTACTTAACTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 906
 Db 50772 GTTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 50827
 QY 907 ATATGTTATGCTGACAGACATCTCAATTGAGTACCTTCAAGTGTCTGAG 962
 Db 50828 GTATTTCACATCTCCACATATTTTGGTTGACACGACCATTTCAAGTGTCTGAG 50883

RESULT 15
 HS227L5 85304 bp DNA linear PRI 22-MAY-2002
 LOCUS Human DNA sequence from clone RP6-227L5 on chromosome Xp11.22-11.3,
 complete sequence.
 ACCESSION AL031585
 VERSION AL031585.2 GI:21212870
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 85304)
 REFERENCE
 AUTHORS Bird,C.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk; clone requests: clonerequest@sanger.ac.uk
 COMMENT On May 25, 2002 this sequence version replaced gi:3980350.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL, SW:
 SWISSPROT, Tr: TrEMBL, Wp: WormPep; Information on the WormPep
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome X, constructed by the Sanger Centre Chromosome X Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chrx
 RP6-227L5 is from the library Rpci-6 constructed by the group of

Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pPAC4.
 FEATURES
 source Location/Qualifiers
 1. 85304
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="RZPD:RZPDP709L05227"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="p11.22-11.3"
 /clone="RP6-227L5"
 /clone_11b="Rpci-6"

ORIGIN
 Query Match 13.1%; Score 261.8; DB 9; Length 85304;
 Best Local Similarity 69.5%; Pred. No. 2.2e-51;
 Matches 448; Conservative 0; Mismatches 177; Indels 20; Gaps 6;
 QY 373 GAGTAGGCTGTAGTAAATATGCTGAGTCTTATGCTATTTCCAGTTTATGAT 432
 Db 81598 GAAGATTTCTTTTGGAGGTTATTAATTAATTTCTAGTAGTAAATTAATTAATG 81657
 QY 433 AGCCACATTAACAGGTAA--AAAGCTGGCGAGTGGCTCACACTGTATCCAG 490
 Db 81658 TTAATTTTAAAAAGTAAATTTTCAAGCTGAGATCTGTGCTCACACTGTATCTCAG 81717
 QY 491 CACTTGGAGGCTGAGGCGAGGAGATCACCTTTGTGAGAGTTGAGACTGAGCTGGC 550
 Db 81718 CATTTGGAGGCTGAGGCGAGGCTGAGATCACTTGAAGTGGGAGTTGAGACAGCTGGC 81777
 QY 551 CAACATGGCGAACTCTGTCTCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 610
 Db 81778 CAACATGGCGAACTCTGTCTCTCTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 81836
 QY 611 GCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCACTTAAACCAAGAGAG 670
 Db 81837 GTGCTGTATCTCAGCTGCTCAGAGGCGGAGAGTGAAGATTTGTTAAACCAAGAGAG 81896
 QY 671 TGGAGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 729
 Db 81897 CAGAGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 81956
 QY 730 CTTTGTCTCAAAAAGAAAAAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 783
 Db 81957 CTCAGTCTCAAAAAGAAAAAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 82016
 QY 784 ACTTAAATACCAATGTATCCCAATATCAATCAATTTCAAGGTAATTAATTAATTAATTAATTA 843
 Db 82017 ACTTAAATACCAATGTATCCCAATATCAATCAATTTCAAGGTAATTAATTAATTAATTA -A 82074
 QY 844 ATTATGATGAGTACTTAACTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 903
 Db 82075 AATATGTTGAATTAATTAATTAATTT-----TTTCAATCAAGTCTTTGAAGTCTGC 82126
 QY 904 TATATATGTTATGCTGACAGACATCTCAATTTGAGTACCTTCAAGTCTTCAAGTCTTCAAGT 963
 Db 82127 TGTATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 82186
 QY 964 AGCCACATGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1008
 Db 82187 AGCCACATGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 82231

Search completed: February 26, 2005, 06:23:59
 Job time: 8539.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 ; Search time 1013.11 seconds
(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacacagagagcccccag.....tcgcagaggggaaacttta 2001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2001 | 100.0 | 11570 | 3 | AAA95905 |
| 2 | 2001 | 100.0 | 11570 | 12 | ADK52482 |
| 3 | 2001 | 100.0 | 11570 | 13 | ADR72623 |
| 4 | 2001 | 100.0 | 11570 | 13 | ADR72875 |
| 5 | 252.6 | 12.6 | 52216 | 6 | AAH28355 |
| 6 | 252.6 | 12.6 | 52216 | 6 | ABL50307 |
| 7 | 252.4 | 12.6 | 56098 | 11 | ACN44922 |
| 8 | 240.2 | 12.0 | 55827 | 8 | ACA60949 |
| 9 | 240.2 | 12.0 | 55827 | 10 | ABX13671 |
| 10 | 240.2 | 12.0 | 58337 | 13 | ADK56454 |
| 11 | 240.2 | 12.0 | 64423 | 13 | ADK56462 |
| 12 | 235.6 | 11.8 | 27189 | 11 | ACN45066 |
| 13 | 229.2 | 11.5 | 110000 | 11 | ACN43998_3 |
| 14 | 227.6 | 11.4 | 2953 | 4 | AAK68359 |
| 15 | 227.6 | 11.4 | 2953 | 4 | AAK68360 |
| 16 | 227.6 | 11.4 | 3377 | 12 | ADK64563 |
| 17 | 227.6 | 11.4 | 13744 | 4 | AAK82406 |
| 18 | 227.6 | 11.4 | 13744 | 4 | AAK82406 |
| 19 | 227.4 | 11.4 | 23071 | 4 | AAK82699 |
| 20 | 227.4 | 11.4 | 23071 | 8 | ADA98598 |

ALIGNMENTS

| RESULT 1 | AAA95905 | standard; DNA; 11570 BP. |
|----------|--|--------------------------|
| ID | AAA95905 | |
| AC | AAA95905; | |
| DT | 02-FEB-2001 | (first entry) |
| XX | | |
| DE | Human KLK-L2 gene. | |
| XX | | |
| KM | Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; | |
| XX | Kallikrein-like protein; serine protease; cytosolic; cancer; | |
| XX | prostate cancer; ds. | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200053776-A2. | |
| XX | | |
| PD | 14-SEP-2000. | |
| XX | | |
| PF | 09-MAR-2000; 2000WO-Ca000258. | |
| XX | | |
| PR | 11-MAR-1999; 99US-0124260P. | |
| XX | 01-APR-1999; 99US-0127386P. | |
| PR | 21-JUL-1999; 99US-0144919P. | |
| XX | | |
| PA | (MOUN) MOUNT SINAI HOSPITAL. | |
| XX | | |
| PI | Yousef GM, Diamandis EP; | |
| XX | WPI; 2000-587440/55. | |
| DR | P-PSDB; AAB21296. | |
| XX | | |
| PT | New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L | |
| XX | protein mediated disorders, especially cancer. | |
| XX | | |
| PS | Claim 1; Page 143-149; 184pp; English. | |
| XX | | |
| CC | The present sequence is the coding sequence of the human KLK-L2 gene, | |
| XX | which encodes a kallikrein-like protein. Kallikreins and kallikrein-like | |
| CC | proteins are a subgroup of the serine protease enzyme family. They | |
| XX | catalyse the selective cleavage of specific polypeptide precursors to | |
| CC | release peptides with potent biological activity. Nucleic acids encoding | |

| | | | | | | | |
|---|----|-------|------|--------|----|------------|---------------------|
| C | 21 | 227.4 | 11.4 | 23071 | 8 | ABX74048 | ABX74048 Human nov |
| | 22 | 227.4 | 11.4 | 23071 | 8 | ADA44323 | Ada44323 Human sec |
| | 23 | 227.4 | 11.4 | 23071 | 10 | ADC20712 | Adc20712 Human sec |
| | 24 | 227.4 | 11.4 | 23071 | 10 | ADK10881 | Adk10881 Human sec |
| | 25 | 226.8 | 11.3 | 12275 | 4 | AAK85734 | Aak85734 Human imm |
| | 26 | 226.2 | 11.3 | 133955 | 11 | ACN45170 | Acn45170 Human gen |
| | 27 | 225.4 | 11.3 | 325791 | 4 | AAK43104 | Aak43104 Human Oes |
| | 28 | 225.2 | 11.3 | 35641 | 6 | ABL64428 | Ab164428 Stomach c |
| | 29 | 225.2 | 11.3 | 35641 | 6 | ABN95727 | Abn95727 Gene #222 |
| | 30 | 224.8 | 11.2 | 629 | 12 | ADM99887 | Adm99887 Human STR |
| | 31 | 224.8 | 11.2 | 629 | 12 | ADM99888 | Adm99888 Human STR |
| | 32 | 223.8 | 11.2 | 15765 | 5 | ABA18276 | Abal18276 Human ner |
| | 33 | 223.8 | 11.2 | 15765 | 5 | ABA19009 | Abal19009 Human ner |
| | 34 | 223.8 | 11.2 | 15765 | 5 | ABA17911 | Abal17911 Human ner |
| | 35 | 223.8 | 11.2 | 15765 | 5 | ABA18250 | Abal18250 Human ner |
| | 36 | 223.6 | 11.2 | 1267 | 4 | AAK85733 | Aak85733 Human imm |
| | 37 | 223.2 | 11.2 | 110000 | 10 | ADG70447_3 | Adg70447 (4 of |
| | 38 | 223.2 | 11.2 | 110000 | 10 | ABZ79565_3 | Abz79565 (4 of |
| | 39 | 222 | 11.1 | 72409 | 13 | ABD33061 | Abd33061 Human can |
| | 40 | 221.6 | 11.1 | 2591 | 4 | AAH16224 | Aah16224 Human CDN |
| | 41 | 221.6 | 11.1 | 2591 | 12 | ADK22885 | Adk22885 Human MP2 |
| | 42 | 221.6 | 11.1 | 2591 | 13 | ADR14082 | Adr14082 Human NF- |
| | 43 | 220.8 | 11.0 | 66685 | 4 | AAK07380 | Aak07380 Human gen |
| | 44 | 220.8 | 11.0 | 66685 | 6 | ABZ73149 | Abz73149 Human CLA |
| | 45 | 220.8 | 11.0 | 96587 | 9 | ADA02984 | Ada02984 Human MAP |

CC kallikrein-like proteins KLK-11, KLK-12, KLK-13, KLK-14, KLK-15 and KLK-16 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No.0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1  GACAAACAAGAGCCCCCAAGAGTGAAGTCCAGGTTCTTCTTGATACGACCCATCTCTCG 60
DB      6000 GACAAACAAGAGCCCCCAAGAGTGAAGTCCAGGTTCTTCTTGATACGACCCATCTCTCG 6055

QY      61  CGGCTTCATCTTCTTCACCTTCTCATTTGTTCTGTTTGAACAGTGCACCTTCCCTAAG 120
DB      6060 CGGCTTCATCTTCTTCACCTTCTCATTTGTTCTGTTTGAACAGTGCACCTTCCCTAAG 6119

QY      121  TCCTCCAGTCTGAATATCAGGTCGTAAGTCAAGAAAGTGCAGAGTCTTACCGA 180
DB      6120 TCCTCCAGTCTGAATATCAGGTCGTAAGTCAAGAAAGTGCAGAGTCTTACCGA 6179

QY      181  GACAGATGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTGCGCAG 240
DB      6180 GACAGATGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTGCGCAG 6239

QY      241  TGAAGACACCTCTCTTTTATTCAGCAGATACACTGAGTGCACCACTCGGTACATGAG 300
DB      6240 TGAAGACACCTCTCTTTTATTCAGCAGATACACTGAGTGCACCACTCGGTACATGAG 6299

QY      301  GTTGCAAAATTCGAGATTCAGCAATTCGCAAGTCAAGAGTCAAGCCCTGTTCTCACA 360
DB      6300 GTTGCAAAATTCGAGATTCGCAAGTTCGCAAGAGTCAAGAGCCCTGTTCTCACA 6359

QY      361  GGTATACCTTGAAGTATGATGTTTATGAAATATATGCTGAGTCTTATATGCTATTC 420
DB      6360 GGTATACCTTGAAGTATGATGTTTATGAAATATATGCTGAGTCTTATATGCTATTC 6419

QY      421  CAGTTTTTATGATGCCACATTAACAGGTAAGGCTGCGGCGACATGAGTCTCACACT 480
DB      6420 CAGTTTTTATGATGCCACATTAACAGGTAAGGCTGCGGCGACATGAGTCTCACACT 6479

QY      481  GTAATCCAGCACTTTGGAGGCTGAGGAGGAGGAGTCACTTTGCTCAGAGTTTGAGA 540
DB      6480 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGGAGTCACTTTGCTCAGAGTTTGAGA 6539

QY      541  CTAGCTGAGCAACATGAGGAAACTGCTCTTAATAAAAAAATATACAAAATTAGCTGGC 600
DB      6540 CTAGCTGAGCAACATGAGGAAACTGCTCTTAATAAAAAAATATAGCTGGC 6599

QY      601  ATGATGCGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCACTTAA 660
DB      6600 ATGATGCGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGGAGACACAAAGATCACTTAA 6659

QY      661  ACCCAGAGAGTGAAGGTGACATGAGTGAATCGGCACTGACCTCAACCTGGAGAGC 720
DB      6660 ACCCAGAGAGTGAAGGTGACATGAGTGAATCGGCACTGACCTCAACCTGGAGAGC 6719

QY      721  AAGATGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG 780
DB      6720 AAGATGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG 6779

QY      781  TTAATCTTAAATAACCAATGATATCCAAATATCAATCATTTCAAGTGAATTAATAA 840
DB      6780 TTAATCTTAAATAACCAATGATATCCAAATATCAATCATTTCAAGTGAATTAATAA 6839

QY      841  ACAATATATGATGATATCTTATCATTTCTTTTCTTTTCAATTAAGTCTTTGAAGT 900
DB      6840 ACAATATATGATGATATCTTATCATTTCTTTTCTTTTCAATTAAGTCTTTGAAGT 6899
```

```
QY      901  GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATAGTACATTTGAGTGCTC 960
DB      6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATAGTACATTTGAGTGCTC 6959

QY      961  AGTAGCCACATGTGGCTAGCACTTACTGTATTTGATGAGCAGGATTTAGAGGAAAGATC 1020
DB      6960 AGTAGCCACATGTGGCTAGCACTTACTGTATTTGATGAGCAGGATTTAGAGGAAAGATC 7019

QY      1021  AAGGCTGTTTTGTATGTTGGCAGGTTGTGACATGATTAAGATTCATATCTAATAG 1080
DB      7020 AAGGCTGTTTTGTATGTTGGCAGGTTGTGACATGATTAAGATTCATATCTAATAG 7079

QY      1081  GGCACCTCCGTGTACAGATGTCAGTTTGGCAGTTTTCAGCGTGTGTAGTTAAGTGC 1140
DB      7080 GGCACCTCCGTGTACAGATGTCAGTTTGGCAGTTTTCAGCGTGTGTAGTTAAGTGC 7139

QY      1141  TTGTTTCAACAAATCTGTATATATGACAGTTTCTTACAGAGTCTGTGTAATCTTGA 1200
DB      7140 TTGTTTCAACAAATCTGTATATATGACAGTTTCTTACAGAGTCTGTGTAATCTTGA 7199

QY      1201  GGAAGAAAAAGAAATCTGTAGGATTTTACAAGAGATATTTAATPACAGGGGATTA 1260
DB      7200 GGAAGAAAAAGAAATCTGTAGGATTTTACAAGAGATATTTAATPACAGGGGATTA 7259

QY      1261  ATTGCAAGCTGCTGGAAGGCTGGAAGAACAAAGTTAAAAAATAAAAACTCTGTGTC 1320
DB      7260 ATTGCAAGCTGCTGGAAGGCTGGAAGAACAAAGTTAAAAAATAAAAACTCTGTGTC 7319

QY      1321  AAGATCTGATTAATAGGCAATTTCAAGAGTGTATTAAGTTAACCCCAATTAATAC 1380
DB      7320 AAGATCTGATTAATAGGCAATTTCAAGAGTGTATTAAGTTAACCCCAATTAATAC 7379

QY      1381  ATGTTTTATGATATGAAACATTAAGGCGCAATATTCAAAAAGTGTGAGGGAGCCTC 1440
DB      7380 ATGTTTTATGATATGAAACATTAAGGCGCAATATTCAAAAAGTGTGAGGGAGCCTC 7439

QY      1441  CTTGAGAGGTGGCATTTGAGCAGAGAAATGATGACACAAAGACCTAATCTGTGAAGT 1500
DB      7440 CTTGAGAGGTGGCATTTGAGCAGAGAAATGATGACACAAAGACCTAATCTGTGAAGT 7499

QY      1501  TTTAAGGGGAAAAAAGGCACTGTCAAGAGCCCTGAGGCGATTAAGAAATTTGGCTGATTC 1560
DB      7500 TTTAAGGGGAAAAAAGGCACTGTCAAGAGCCCTGAGGCGATTAAGAAATTTGGCTGATTC 7559

QY      1561  AAAAGAAAGAGAAACCAATGCAATGAGAAACAAAGTGGGGGCAACATGAGAAAGTG 1620
DB      7560 AAAAGAAAGAGAAACCAATGCAATGAGAAACAAAGTGGGGGCAACATGAGAAAGTG 7619

QY      1621  ACGCTGAGAGTGTAGGCAAGGGGCGAATGCTGTCAAGTATTTCTTGTCACCAACAG 1680
DB      7620 ACGCTGAGAGTGTAGGCAAGGGGCGAATGCTGTCAAGTATTTCTTGTCACCAACAG 7679

QY      1681  GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTGAGAAAGCAGAAATTTAAATCAAC 1740
DB      7680 GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTGAGAAAGCAGAAATTTAAATCAAC 7739

QY      1741  TGTTCATCAACACAGACCCCTTCTGTATTCAGGCTCCCAAGGGATCTAGAAAGAGC 1800
DB      7740 TGTTCATCAACACAGACCCCTTCTGTATTCAGGCTCCCAAGGGATCTAGAAAGAGC 7799

QY      1801  AGTTAACAGCTCTCATTAAGCAGGGTGTGTGTTTCAACAGTATAGAGTGTGGGATTT 1860
DB      7800 AGTTAACAGCTCTCATTAAGCAGGGTGTGTGTTTCAACAGTATAGAGTGTGGGATTT 7859

QY      1861  CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCCCAATTTGTACAACTGTACAA 1920
DB      7860 CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCCCAATTTGTACAACTGTACAA 7919

QY      1921  ATGCAACCTCCGAGCTCCCATATTTCTCTGAGGCTCTGGGTGTGGGGGGGTGGG 1980
DB      7920 ATGCAACCTCCGAGCTCCCATATTTCTCTGAGGCTCTGGGTGTGGGGGGGTGGG 7979

QY      1981  TTGCGAGGGGGAAAACTTTTA 2001
```

Db 7980 TTGCGAGGGGAAACTTTTA 8000

RESULT 2
ADK52482
ID ADK52482 standard; DNA; 11570 BP.

XX ADK52482;

DT 03-JUN-2004 (first entry)

XX Human kallikrein 5 encoding sequence.

XX kallikrein 5; cancer; Cytostatic; Immunostimulant; db.

OS Homo sapiens.

PN MO2004021008-A2.

XX 11-MAR-2004.

PF 28-AUG-2003; 2003MO-CA001310.

XX 28-AUG-2002; 2002US-0407333P.

PA (MOUN) MOUNT SINAI HOSPITAL.

PI (YOUS/) YOUSEF G.

XX Diamandis BP;

DR MPI: 2004-239232/22.

XX P-PSDB; ADK52481.

PT Detecting kallikrein 5 associated with cancer, useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma, PT comprises comparing the detected amount of kallikrein 5 in a sample with a standard sample.

XX Disclosure; SEQ ID NO 2; 59pp; English.

XX The present invention relates to detecting kallikrein 5 associated with breast or ovarian cancer in a patient comprises detecting in the sample kallikrein 5 and comparing the detected amount with an amount detected for a standard. The method is useful in detecting kallikrein 5 associated with breast or ovarian cancer in a patient. The methods and kits are useful in diagnosing, monitoring, detecting, imaging and treating breast or ovarian carcinoma. The kallikrein 5 is useful in preparing a vaccine for preventing and treating breast and ovarian cancer and for stimulating or enhancing antibody production or for inducing an immune response. The present sequence represents human kallikrein 5 encoding sequence.

XX Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 12; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAACCAAGAGCCCAAGAGTGTGTCAGGTTCTTCTTGAATACCAACCAATCTCTGC 60

Db 6000 GACAAACCAAGAGCCCAAGAGTGTGTCAGGTTCTTCTTGAATACCAACCAATCTCTGC 6059

QY 61 CGGCTTCATCTTTCACATTTCTCATTTGTTCTGTTTGAAGTGCATCTCCCTAAG 120

Db 6060 CGGCTTCATCTTTCACATTTCTCATTTGTTCTGTTTGAAGTGCATCTCCCTAAG 6119

QY 121 TCCTCAGTGTGAATATCAGGTGCTAAGTCAAGAAAGTGGAGATCTTACCCGA 180

Db 6120 TCCTCAGTGTGAATATCAGGTGCTAAGTCAAGAAAGTGGAGATCTTACCCGA 6179

QY 181 GACAGATAGATGACACCATGTTCTGCGCCGCTGCAAAAGCAGGTAGAGACTCTGCGAG 240

Db 6180 GACAGATAGATGACACCATGTTCTGCGCCGCTGCAAAAGCAGGTAGAGACTCTGCGAG 6239

QY 241 TGAGACACCTCTCTTATTCAGCAGATACACATGAGTGCCAACTCGGTAACTGAGC 300

Db 6240 TGAGACACCTCTCTTATTCAGCAGATACACATGAGTGCCAACTCGGTAACTGAGC 6299

QY 301 GTTGCCAAATTCAGAAATCCAGCAATTCGCAAGACGTCAAGACCCTGTTCTCAGA 360

Db 6300 GTTGCCAAATTCAGAAATTCAGCAATTCGCAAGACGTCAAGACCCTGTTCTCAGA 6359

QY 361 GCTCATACCCCTAGAGTGTGTTTGAAGAAATATGCTGAGCTGCTTATGCTATTC 420

Db 6360 GCTCATACCCCTAGAGTGTGTTTGAAGAAATATGCTGAGCTGCTTATGCTATTC 6419

QY 421 CAGTTTTTAAAGTCAATTAATAAGGTAATAAGGTCGCGGAGGCTCACACCT 480

Db 6420 CAGTTTTTAAAGTCAATTAATAAGGTAATAAGGTCGCGGAGGCTCACACCT 6479

QY 481 GTAATCCAGACCTTGGAGGCTGAGGAGGAGATCACTTTGGTCAAGAGTTTGA 540

Db 6480 GTAATCCAGACCTTGGAGGCTGAGGAGGAGATCACTTTGGTCAAGAGTTTGA 6539

QY 541 CTAGCTGGCCAAATGCGGAAACTGTCTCTTAAAAAAATACAAAATTTAGCCTG 600

Db 6540 CTAGCTGGCCAAATGCGGAAACTGTCTCTTAAAAAAATACAAAATTTAGCCTG 6599

QY 601 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACCAAGAACTTAA 660

Db 6600 ATGCTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACCAAGAACTTAA 6659

QY 661 ACCCAGAGGTGAGGTGCTCAGTCAAGTCAAGTCTGCTCACTCACTCACTGAGAG 720

Db 6660 ACCCAGAGGTGAGGTGCTCAGTCAAGTCAAGTCTGCTCACTCACTCACTGAGAG 6719

QY 721 AGAGTCACTTTTGTCTCAAAAAAGAAAAAAGTAAAAAAGGAGAG 780

Db 6720 AGAGTCACTTTTGTCTCAAAAAAGAAAAAAGTAAAAAAGGAGAG 6779

QY 781 TTAACCTTAATAACCAATGTAATCCCAATATCAATTAATTAATTAATAA 840

Db 6780 TTAACCTTAATAACCAATGTAATCCCAATATCAATTAATTAATTAATAA 6839

QY 841 ACAATATGATGATGATTAATCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 900

Db 6840 ACAATATGATGATGATTAATCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6899

QY 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATGCTACATTTCAAGTCTC 960

Db 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATGCTACATTTCAAGTCTC 6959

QY 961 AGTAGCCATGAGTGTGCTACAGTACTGTAATGGAATGCGACCGAATCTAAGAGAAAGATC 1020

Db 6960 AGTAGCCATGAGTGTGCTACAGTACTGTAATGGAATGCGACCGAATCTAAGAGAAAGATC 7019

QY 1021 AGGCTGTTTTGTAATGTTGGGAGGTTGTCATGATTAAGATATCAATATTAATAG 1080

Db 7020 AGGCTGTTTTGTAATGTTGGGAGGTTGTCATGATTAAGATATCAATATTAATAG 7079

QY 7080 GGCACCTCGGTGTAAGATGTCAGTTTGGCACTTTCAAGGCTGTGTATTAAGTGC 1140

Db 1081 GGCACCTCGGTGTAAGATGTCAGTTTGGCACTTTCAAGGCTGTGTATTAAGTGC 7139

QY 1141 TTGTTTCAACAAAATCTGTAATATGACATTTTCTGCAAGTGTGTAATAATCTTGA 1200

Db 7140 TTGTTTCAACAAAATCTGTAATATGACATTTTCTGCAAGTGTGTAATAATCTTGA 7199

QY 1201 GGAAGAAAAGAAATCTGTATGATTTTAAAGAGAAATTTAATACAGGGAGTTA 1260

Db 7200 GGAAGAAAAGAAATCTGTATGATTTTAAAGAGAAATTTAATACAGGGAGTTA 7259

QY 1261 ATTGCAAGCTGTGGAAGGCTGTGAGAACAAAGTTAAAAAATTAATAAATCTGTGCTC 1320

Db 7260 ATTGCAAGCTGTGGAAGGCTGTGAGAACAAAGTTAAAAAATTAATAAATCTGTGCTC 7319

| | | | |
|----------|-------------|---|------|
| OY | 1321 | AAGATCTGCATTAATAGGGCAATTTCAGAGTGTAAAGTTAAACCCCAAAATPAAAC | 1380 |
| Db | 7320 | AAGATCTGCATTAANTAGGGCAATTTCAAGATGTGTAAAGTTTAACTTAACTCCCAATPAAAC | 7379 |
| OY | 1381 | ATGGTTTTAGATAGTAACAATAAGGCCAATATTCAAAGAGTGTCAAGGGAGCCTC | 1440 |
| Db | 7380 | ATGGTTTTAGATAGTAACAATAAGGCCCAATATTCAAAGAGTGTCAAGGGAGCCTC | 7439 |
| OY | 1441 | CTTGGAGAGGTGGCAATTTGACACAGAAATGGAATGACACAAAGAGCTTAACTCGTAAGT | 1500 |
| Db | 7440 | CTTGGAGAGGTGGCAATTTGACACAGAAATGGAATGACACAAAGAGCTTAACTCGTAAGT | 7499 |
| OY | 1501 | TTAAGGGGAAAGAAAAGGCACTGTCAAAAGGCCCTGAGGCACTGAAGAAATTTGGCTGAATC | 1560 |
| Db | 7500 | TTAAGGGGAAAGAAAAGGCACTGTCAAAAGGCCCTGAGGCACTGAAGAAATTTGGCTGAATC | 7559 |
| OY | 1561 | AAAGAGAGAGAGAAACCAATGCAACTGGAGAACAAAGTGGGGGCAACAGTAGAAAGTG | 1620 |
| Db | 7560 | AAAGAGAGAGAGAGAAACCAATGCAACTGGAGAACAAAGTGGGGGCAACAGTAGAAAGTG | 7619 |
| OY | 1621 | ACGCTGGAGGTGTAGGACGAGGGGCGAATGCTCTGCAAGTATTTCTTGGTCACCAACACGA | 1680 |
| Db | 7620 | ACGCTGGAGGTGTAGGACGAGGGGCGAATGCTCTGCAAGTATTTCTTGGTCACCAACACGA | 7679 |
| OY | 1681 | GCTTCCCCTATGTTCTTAATGGAAGCTATCTGTTGAGGAGAACAGAAATTTAAATCAAC | 1740 |
| Db | 7680 | GCTTCCCCTATGTTCTTAATGGAAGCTATCTGTTGAGGAGAACAGAAATTTAAATCAAC | 7739 |
| OY | 1741 | TGTTACATCAACACGACACCTTCTCTGTATTCAGGCTCCCAAGGATCTGAAGAGACGTA | 1800 |
| Db | 7740 | TGTTACATCAACACGACACCTTCTCTGTATTCAGGCTCCCAAGGATCTGAAGAGACGTA | 7799 |
| OY | 1801 | AGTTAACAGCTCTCATTTAGAGAGGTGTGTGTTTCAACAGTAGTTAGAGAGCTGGGATTT | 1860 |
| Db | 7800 | AGTTAACAGCTCTCATTTAGAGAGGTGTGTGTTTCAACAGTAGTTAGAGAGCTGGGATTT | 7859 |
| OY | 1861 | CAGAGTATCTCCAGTCCCAATGCTATGAAAGCTCCGCCCAATTTGTAACAACCTGACAA | 1920 |
| Db | 7860 | CAGAGTATCTCCAGTCCCAATGCTATGAAAGCTCCGCCCAATTTGTAACAACCTGACAA | 7919 |
| OY | 1921 | ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTCTGTGACCCTGGAGTGTGGGGGGGTGGG | 1980 |
| Db | 7920 | ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTCTGTGACCCTGGAGTGTGGGGGGGTGGG | 7979 |
| OY | 1981 | TTGCGAGGGGGAATTTTAA 2001 | |
| Db | 7980 | TTGCGAGGGGGAATTTTAA 8000 | |
| RESULT 3 | | | |
| ADR72623 | | | |
| ID | ADR72623 | standard; DNA; 11570 BP. | |
| XX | ADR72623; | | |
| AC | 02-DEC-2004 | (first entry) | |
| XX | | | |
| DT | | | |
| XX | | | |
| DE | | Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2. | |
| XX | | | |
| KW | | kallikrein 5; renal cell carcinoma; cytostatic; tumour growth; | |
| KW | | cancer metastasis; chemotherapy; human; serine protease; | |
| XX | | chromosome 19q13.4; KLK5; ds; gene. | |
| XX | | | |
| OS | | Homo sapiens. | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 2221..11247 | |
| FT | | /*tag= b | |
| FT | | /product= "Human renal cell carcinoma-related kallikrein | |
| FT | exon | 5 (hK5) protein" | |
| FT | | 2221..2293 | |
| FT | | /*tag= a | |
| FT | | /number= 1 | |

| FT | intron | 2294. .4761 | /*cag= c | /number= 1 | | | | | |
|----|--------|---------------|-----------|-------------|--------|-------------|---------------|------------|-------------|
| FT | exon | 4762. .5023 | /*cag= d | FT | intron | 5024. .5762 | /*cag= e | /number= 2 | |
| FT | exon | 5763. .6019 | /*cag= f | /number= 3 | FT | intron | 6020. .6104 | /*cag= g | /number= 4 |
| FT | exon | 6105. .6238 | /*cag= h | /number= 5 | FT | exon | 6239. .11091 | /*cag= i | /number= 6 |
| FT | exon | 11092. .11247 | /*cag= j | /number= 7 | FT | exon | 11248. .11403 | /*cag= k | /number= 8 |
| FT | exon | 11404. .11559 | /*cag= l | /number= 9 | FT | exon | 11560. .11715 | /*cag= m | /number= 10 |
| FT | exon | 11716. .11871 | /*cag= n | /number= 11 | FT | exon | 11872. .12027 | /*cag= o | /number= 12 |
| FT | exon | 12028. .12183 | /*cag= p | /number= 13 | FT | exon | 12184. .12339 | /*cag= q | /number= 14 |
| FT | exon | 12340. .12495 | /*cag= r | /number= 15 | FT | exon | 12496. .12651 | /*cag= s | /number= 16 |
| FT | exon | 12652. .12807 | /*cag= t | /number= 17 | FT | exon | 12808. .12963 | /*cag= u | /number= 18 |
| FT | exon | 12964. .13119 | /*cag= v | /number= 19 | FT | exon | 13120. .13275 | /*cag= w | /number= 20 |
| FT | exon | 13276. .13431 | /*cag= x | /number= 21 | FT | exon | 13432. .13587 | /*cag= y | /number= 22 |
| FT | exon | 13588. .13743 | /*cag= z | /number= 23 | FT | exon | 13744. .13899 | /*cag= aa | /number= 24 |
| FT | exon | 13900. .14055 | /*cag= ab | /number= 25 | FT | exon | 14056. .14211 | /*cag= ac | /number= 26 |
| FT | exon | 14212. .14367 | /*cag= ad | /number= 27 | FT | exon | 14368. .14523 | /*cag= ae | /number= 28 |
| FT | exon | 14524. .14679 | /*cag= af | /number= 29 | FT | exon | 14680. .14835 | /*cag= ag | /number= 30 |
| FT | exon | 14836. .14991 | /*cag= ah | /number= 31 | FT | exon | 14992. .15147 | /*cag= ai | /number= 32 |
| FT | exon | 15148. .15303 | /*cag= aj | /number= 33 | FT | exon | 15304. .15459 | /*cag= ak | /number= 34 |
| FT | exon | 15460. .15615 | /*cag= al | /number= 35 | FT | exon | 15616. .15771 | /*cag= am | /number= 36 |
| FT | exon | 15772. .15927 | /*cag= an | /number= 37 | FT | exon | 15928. .16083 | /*cag= ao | /number= 38 |
| FT | exon | 16084. .16239 | /*cag= ap | /number= 39 | FT | exon | 16240. .16395 | /*cag= aq | /number= 40 |
| FT | exon | 16396. .16551 | /*cag= ar | /number= 41 | FT | exon | 16552. .16707 | /*cag= as | /number= 42 |
| FT | exon | 16708. .16863 | /*cag= at | /number= 43 | FT | exon | 16864. .17019 | /*cag= au | /number= 44 |
| FT | exon | 17020. .17175 | /*cag= av | /number= 45 | FT | exon | 17176. .17331 | /*cag= aw | /number= 46 |
| FT | exon | 17332. .17487 | /*cag= ax | /number= 47 | FT | exon | 17488. .17643 | /*cag= ay | /number= 48 |
| FT | exon | 17644. .17799 | /*cag= az | /number= 49 | FT | exon | 17800. .17955 | /*cag= ba | /number= 50 |
| FT | exon | 17956. .18111 | /*cag= bb | /number= 51 | FT | exon | 18112. .18267 | /*cag= bc | /number= 52 |
| FT | exon | 18268. .18423 | /*cag= bd | /number= 53 | FT | exon | 18424. .18579 | /*cag= be | /number= 54 |
| FT | exon | 18580. .18735 | /*cag= bf | /number= 55 | FT | exon | 18736. .18891 | /*cag= bg | /number= 56 |
| FT | exon | 18892. .19047 | /*cag= bh | /number= 57 | FT | exon | 19048. .19203 | /*cag= bi | /number= 58 |
| FT | exon | 19204. .19359 | /*cag= bj | /number= 59 | FT | exon | 19360. .19515 | /*cag= bk | /number= 60 |
| FT | exon | 19516. .19671 | /*cag= bl | /number= 61 | FT | exon | 19672. .19827 | /*cag= bm | /number= 62 |
| FT | exon | 19828. .19983 | /*cag= bn | /number= 63 | FT | exon | 19984. .20139 | /*cag= bo | /number= 64 |
| FT | exon | 20140. .20295 | /*cag= bp | /number= 65 | FT | exon | 20296. .20451 | /*cag= bq | /number= 66 |
| FT | exon | 20452. .20607 | /*cag= br | /number= 67 | FT | exon | 20608. .20763 | /*cag= bs | /number= 68 |
| FT | exon | 20764. .20919 | /*cag= bt | /number= 69 | FT | exon | 20920. .21075 | /*cag= bu | /number= 70 |
| FT | exon | 21076. .21231 | /*cag= bv | /number= 71 | FT | exon | 21232. .21387 | /*cag= bw | /number= 72 |
| FT | exon | 21388. .21543 | /*cag= bx | /number= 73 | FT | exon | 21544. .21699 | /*cag= by | /number= 74 |
| FT | exon | 21700. .21855 | /*cag= bz | /number= 75 | FT | exon | 21856. .22011 | /*cag= ca | /number= 76 |
| FT | exon | 22012. .22 | | | | | | | |

| | | | |
|----|------|---|------|
| QY | 1 | TACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTCCTTCTTGAATACGACCACTCTCGC | 60 |
| Db | 6000 | GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGGTCCTTCTTGAATACGACCACTCTCGC | 6059 |
| QY | 61 | CGCCTTCACATCTTTCTCCACATTCCTGATGTGTTCCCTGTTTGAACAGTGCACCTTCCCTAAG | 120 |
| Db | 6060 | CGCCTTCACATCTTTCTCCACATTCCTGATGTGTTCCCTGTTTGAACAGTGCACCTTCCCTAAG | 6119 |
| QY | 121 | TCCTCCAGTGCCTTGAATATACAGCGTGTACAGTGCAGAAAAGTGCAGAGATCCTTACCCGA | 180 |
| Db | 6120 | TCCTCCAGTGCCTTGAATATACAGCGTGTACAGTGCAGAAAAGTGCAGAGATCCTTACCCGA | 6179 |
| QY | 181 | GACAGATATGATGACACCAATGTTCTGTCGGCCGGTGAACAAAGCAGGTATGACATCTCCGACAG | 240 |
| Db | 6180 | GACAGATATGATGACACCAATGTTCTGTCGGCCGGTGAACAAAGCAGGTATGACATCTCCGACAG | 6239 |
| QY | 241 | TGAGGACACCTCTCTTTATTCAGCAGATACACATCGAGTGCACATCGGTAACTGAGC | 300 |
| Db | 6240 | TGAGGACACCTCTCTTTATTCAGCAGATATACACATCGAGTGCACATCGGTAACTGAGC | 6299 |
| QY | 301 | GTTCGCAAAATCTGAGAAATCCAGCAATTCGCAAGCAGTCAAGAACCCCTGTTCTCACAGA | 360 |
| Db | 6300 | GTTCGCAAAATCTGAGAAATCCAGCAATTCGCAAGCAGTCAAGAACCCCTGTTCTCACAGA | 6359 |
| QY | 361 | GCTCATACCCCTAGAGTATGAGTGTGTTTATGAGAAATATGCTGAGTGTCTTATGTCAATTC | 420 |
| Db | 6360 | GCTCATACCCCTAGAGTATGAGTGTGTTTATGAGAAATATGCTGAGTGTCTTATGTCAATTC | 6419 |
| QY | 421 | CAGTTTTTTTATGATGACCATTTAAACAGGTAAAAAAGGCTGGGCGCAGTGGCTCACACT | 480 |
| Db | 6420 | CAGTTTTTTTATGATGACCATTTAAACAGGTAAAAAAGGCTGGGCGCAGTGGCTCACACT | 6479 |
| QY | 481 | GTAATCCAGCATTTTGGAGGCTGAGGCAAGCAGATCACTTTGGTCAGAGAGTTTGAGA | 540 |
| Db | 6480 | GTAATCCAGCATTTTGGAGGCTGAGGCAAGCAGATCACTTTGGTCAGAGAGTTTGAGA | 6539 |
| QY | 541 | CTAGCCTGGCCAAACATGCGCAAACTCTGCTCTTAAAAAAAATACAAAATTAAGCTGGC | 600 |
| Db | 6540 | CTAGCCTGGCCAAACATGCGCAAACTCTGCTCTTAAAAAAAATACAAAATTAAGCTGGC | 6599 |
| QY | 601 | ATGTGTGGCGGGGCGCTGTATCTCAGCTGCTCAGGAGGCGAGACACAAAGTACACTTAA | 660 |
| Db | 6600 | ATGTGTGGCGGGGCGCTGTATCTCAGCTGCTCAGGAGGCGAGACACAAAGTACACTTAA | 6659 |
| QY | 661 | ACCAGAGGTGAGGTTCGACGTGAGTCTGTGCACCTCACTCCAACTGCGAGAC | 720 |
| Db | 6660 | ACCAGAGGTGAGGTTCGACGTGAGTCTGTGCACCTCACTCCAACTGCGAGAC | 6719 |
| QY | 721 | AGAGTGCACATTTTGTCTCAAAAAGAAAAAACAATGTAATAAAGAAACAGGTGAAG | 780 |
| Db | 6720 | AGAGTGCACATTTTGTCTCAAAAAGAAAAAACAATGTAATAAAGAAACAGGTGAAG | 6779 |
| QY | 781 | TTAACTTTAATPAAACCAATGTATCCCAANTCAATCATTTCAAAGTGTAAATTAATATAA | 840 |
| Db | 6780 | TTAACTTTAATPAAACCAATGTATCCCAANTCAATCATTTCAAAGTGTAAATTAATATAA | 6839 |
| QY | 841 | ACAATTAATGAATGAGATACCTTACATCTTTCTGTTTCATATTAATTAAGCTTTGAAGT | 900 |
| Db | 6840 | ACAATTAATGAATGAGATACCTTACATCTTTCTGTTTCATATTAATTAAGCTTTGAAGT | 6899 |
| QY | 901 | GAGTATATATGTTATGCTGACAGCACATCTCAATTTGGAATGACTCAATTCAGGTGCTC | 960 |
| Db | 6900 | GAGTATATATGTTATGCTGACAGCACATCTCAATTTGGAATGACTCAATTCAGGTGCTC | 6959 |
| QY | 961 | AGTACCCACATGTGGCTGACAGTTATCTGTATTTGATGTCGACGAGATGAGAGGAAAGATC | 1020 |
| Db | 6960 | AGTACCCACATGTGGCTGACAGTTATCTGTATTTGATGTCGACGAGATGAGAGGAAAGATC | 7019 |
| QY | 1021 | AGGGCTGTTTGTATGTTGTTGGGCAAGTTGTGTGCATGCAATTAAGATACATATCTTAATAG | 1080 |
| Db | 7020 | AGGGCTGTTTGTATGTTGTTGGGCAAGTTGTGTGCATGCAATTAAGATACATATCTTAATAG | 7079 |
| QY | 1081 | GGCACTCCGTGTACAGATGTCAAGTTTGGCAGTTTCAAGCGTGTGTGATTAATGTCTC | 1140 |

| | | | |
|----|------|--|------|
| Db | 7080 | GGCACTCCGCTTACAAATGCAATGCACTTTTGGCAGTTTTCAGGCCGTGTGTAGTTAAAGTCTC | 7139 |
| Oy | 1141 | TTGTTTCAACAAAATCTGTAAATATGACAGTTTTCAGCAAGTGTCTGTAAATAATCTTTGA | 1200 |
| Db | 7140 | TTGTTTCAACAAAATCTGTAAATATGACAGTTTTCAGCAAGTGTCTGTAAATAATCTTTGA | 7199 |
| Oy | 1201 | GGAAAGAAAAGAAATCTGTGTAGTATTTTTCACAGAAATATTTAAATACAGGGGATTA | 1260 |
| Db | 7200 | GGAAAGAAAAGAAATCTGTGTAGTATTTTTCACAGAAATATTTAAATACAGGGGATTA | 7259 |
| Oy | 1261 | ATTGCAAAAGCTGCTGAAAGGGCTGGAGGAACTAAAGTTAAATAATAAAAACTCTGTGTCTC | 1320 |
| Db | 7260 | ATTGCAAAAGCTGCTGAAAGGGCTGGAGGAACTAAAGTTAAATAATAAAAACTCTGTGTCTC | 7319 |
| Oy | 1321 | AAGAACTGTGCAATAAATAGGGCAATTTTCAGAGAGTGTAAAGTTTAACTCCCAATAATAAAC | 1380 |
| Db | 7320 | AAGAACTGTGCAATAAATAGGGCAATTTTCAGAGAGTGTAAAGTTTAACTCCCAATAATAAAC | 7379 |
| Oy | 1381 | ATGCTTTTAAAGTATGTAACAAATTAAGGCCCAATATTCAAAAGAGTGTCTCAGGGAGCTTC | 1440 |
| Db | 7380 | ATGCTTTTAAAGTATGTAACAAATTAAGGCCCAATATTCAAAAGAGTGTCTCAGGGAGCTTC | 7439 |
| Oy | 1441 | CTTGGAAGAGGTGGCAATTTGAGCAGAGAAATGGAATACAAAGAAAGCTTAACTCGTGAAGT | 1500 |
| Db | 7440 | CTTGGAAGAGGTGGCAATTTGAGCAGAGAAATGGAATACAAAGAAAGCTTAACTCGTGAAGT | 7499 |
| Oy | 1501 | TTTAAAGGGAAAAGAAAAGGCACTGTCAAAAGGCCCTGAGAGCAGTAAGAAATTTTGCTGATTC | 1560 |
| Db | 7500 | TTTAAAGGGAAAAGAAAAGGCACTGTCAAAAGGCCCTGAGAGCAGTAAGAAATTTTGCTGATTC | 7559 |
| Oy | 1561 | AAAGAAGAAGAGAAACCAATGCACTTGAGAAACAAAAGTGGGGGCAACAGTAGAAAGTG | 1620 |
| Db | 7560 | AAAGAAGAAGAGAAACCAATGCACTTGAGAAACAAAAGTGGGGGCAACAGTAGAAAGTG | 7619 |
| Oy | 1621 | ACGCTGAAGGTATGAGAGGGGGCAAGCTCTGCAATATTTCTTGATCAACACACGA | 1680 |
| Db | 7620 | ACGCTGAAGGTATGAGAGGGGGCAAGCTCTGCAATATTTCTTGATCAACACACGA | 7679 |
| Oy | 1681 | GCTTCCCTATGTTCTTAATGAGAGCTGTATCTGTTTGAAGAAAGACGAATTTTAAATCAAAC | 1740 |
| Db | 7680 | GCTTCCCTATGTTCTTAATGAGAGCTGTATCTGTTTGAAGAAAGACGAATTTTAAATCAAAC | 7739 |
| Oy | 1741 | TGTTTACATCAACACGACCCCTTCTCTGTATTCAGGCTCCCAAGGAATCTAAGAGACGTA | 1800 |
| Db | 7740 | TGTTTACATCAACACGACCCCTTCTCTGTATTCAGGCTCCCAAGGAATCTAAGAGACGTA | 7799 |
| Oy | 1801 | AGTTTAAACAGCTCTCATTAAGCAGGGGTGTGTTCACACATAGTTAAGAAAGCTGGGAAAT | 1860 |
| Db | 7800 | AGTTTAAACAGCTCTCATTAAGCAGGGGTGTGTTCACACATAGTTAAGAAAGCTGGGAAAT | 7859 |
| Oy | 1861 | CAGGAGTACTCCAGTCCCATGAGCTATGATAAAAGCTCCCCCAAAATTTGTACAAACTGTGCAA | 1920 |
| Db | 7860 | CAGGAGTACTCCAGTCCCATGAGCTATGATAAAAGCTCCCCCAAAATTTGTACAAACTGTGCAA | 7919 |
| Oy | 1921 | ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCCCTGGAGTGTGGGGGGGTGGG | 1980 |
| Db | 7920 | ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCCCTGGAGTGTGGGGGGGTGGG | 7979 |
| Oy | 1981 | TTTGCAAGGGGGAAAACCTTTTA 2001 | |
| Db | 7980 | TTTGCAAGGGGGAAAACCTTTTA 8000 | |

| | | |
|----|---|---|
| XX | | kallikrein 5; tumour marker; ovarian cancer; |
| KW | | epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4; |
| KM | | KLK5; ds; gene. |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 2221..11247 |
| FT | | /tag= b |
| FT | | /product= "human ovarian cancer-related tumour marker |
| FT | exon | kallikrein 5 (hK5) protein" |
| FT | | 2221..22293 |
| FT | | /tag= a |
| FT | | /number= 1 |
| FT | intron | 2294..4761 |
| FT | | /tag= c |
| FT | | /number= 1 |
| FT | exon | 4762..5023 |
| FT | | /tag= d |
| FT | | /number= 2 |
| FT | intron | 5024..5762 |
| FT | | /tag= e |
| FT | | /number= 2 |
| FT | exon | 5763..6019 |
| FT | | /tag= f |
| FT | | /number= 3 |
| FT | intron | 6020..6104 |
| FT | | /tag= g |
| FT | | /number= 3 |
| FT | exon | 6105..6238 |
| FT | | /tag= h |
| FT | | /number= 4 |
| FT | exon | 6239..11091 |
| FT | | /tag= i |
| FT | | /number= 4 |
| FT | exon | 11092..11247 |
| FT | | /tag= j |
| FT | | /number= 5 |
| XX | | |
| PN | WO2004075713-A2. | |
| XX | | |
| PD | 10-SEP-2004. | |
| XX | | |
| PF | 26-FEB-2004; 2004WO-CA000281. | |
| XX | | |
| PR | 26-FEB-2003; 2003US-0450406P. | |
| XX | | |
| PA | (MOUN) MOUNT SINAI HOSPITAL. | |
| XX | | |
| PI | Diamandis EP; | |
| XX | | |
| XX | WPI; 2004-661815/64. | |
| DR | P-PSDB; ADR72873. | |
| XX | | |
| PT | Kallikrein markers detection method for detecting ovarian cancer in | |
| PT | patient, involves detecting kallikrein markers and CA125 in sample | |
| PT | obtained from patient, and comparing detected amounts with standard | |
| PT | amounts. | |
| XX | | |
| PS | Example 2; SEQ ID NO 5; 102pp; English. | |
| CC | | |
| CC | The invention relates to a novel method for detecting a plurality of | |
| CC | kallikrein markers associated with ovarian cancer. The method comprises | |
| CC | obtaining a sample from a patient and detecting in the sample a plurality | |
| CC | of kallikrein markers, and optionally carbohydrate antigen CA125, wherein | |
| CC | the kallikrein markers are selected from the group consisting of | |
| CC | kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and | |
| CC | kallikrein 11. The detected amounts of the kallikrein markers are | |
| CC | compared with standard amounts. The method of the invention may be useful | |
| CC | for detecting kallikrein markers associated with ovarian cancer in a | |
| CC | patient and thus for detecting ovarian cancer, particularly epithelial | |
| CC | ovarian carcinoma. The current sequence is that of the human ovarian | |

| CC | Cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention |
|----------------------------|--|
| CC | which encodes a secreted serine protease and is located at chromosome |
| CC | 19q13.4. |
| XX | |
| XX | Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other; |
| Query Match | 100.0%; Score 2001; DB 13; Length 11570; |
| Best Local Similarity | 100.0%; Pred. No. 0; |
| Matches 2001; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | 1 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGTTCCTTTGATACCGACCCATCTCTGC 60 |
| DB | 6000 GACAAACCAAGAGCCCCCAAGGTGAGTGTCCAGTTCCTTTGATACCGACCCATCTCTGC 6059 |
| QY | 61 CGCCCTTCATCTTTCTCCACTTCATTCATTCGTTTCCTGTTTGAAGTGCACCTTCCCTAAG 120 |
| DB | 6060 CGCCCTTCATCTTTCTCCACTTCATTCATTCGTTTCCTGTTTGAAGTGCACCTTCCCTAAG 6119 |
| QY | 121 TCCTCCAGTGTCTTAATATACAGCGTGTAAATCAAGAAAGGTGCGAGATCTTATCCCGA 180 |
| DB | 6120 TCCTCCAGTGTCTTAATATACAGCGTGTAAATCAAGAAAGGTGCGAGATCTTATCCCGA 6179 |
| QY | 181 GACGATATGATGAAACCATGTTCTTGCGCCCGGTGCAAGAAGATGATGAGATCTCCGCGAG 240 |
| DB | 6180 GACGATATGATGAAACCATGTTCTTGCGCCCGGTGCAAGAAGATGATGAGATCTCCGCGAG 6239 |
| QY | 241 TGAGAGACACCTCTCTTATTCAGCAGATATACACTGATGTCCAACTCGTAACTAGAGAGC 300 |
| DB | 6240 TGAGAGACACCTCTCTTATTCAGCAGATATACACTGATGTCCAACTCGTAACTAGAGAGC 6299 |
| QY | 301 GTTGCCAAATTCGAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGAGA 360 |
| DB | 6300 GTTGCCAAATTCGAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGAGA 6359 |
| QY | 361 GCTCATPCCCTAGAGTGTGTGTGTTTGTAGTAAATATATGCTGAGCTCTTATATGCACTTC 420 |
| DB | 6360 GCTCATPCCCTAGAGTGTGTGTGTTTGTAGTAAATATATGCTGAGCTCTTATATGCACTTC 6419 |
| QY | 421 CAGTTTTTATAGTACCAATTAAACAGGTAAACAAAGGCTGGCGCAGTGGCTCAACCT 480 |
| DB | 6420 CAGTTTTTATAGTACCAATTAAACAGGTAAACAAAGGCTGGCGCAGTGGCTCAACCT 6479 |
| QY | 481 GTAATTCACAGACTTTTGGGAGGCTGAGGCGAGGCGAGATCACTTTGGTCAGAGTTTGA 540 |
| DB | 6480 GTAATTCACAGACTTTTGGGAGGCTGAGGCGAGGCGAGATCACTTTGGTCAGAGTTTGA 6539 |
| QY | 541 CTACGCTGGCCAAATGAGCGAAATCTGTCTCTTAAACAAATTAACAATTAAGCTGGC 600 |
| DB | 6540 CTACGCTGGCCAAATGAGCGAAATCTGTCTCTTAAACAAATTAACAATTAAGCTGGC 6599 |
| QY | 601 ATGTTGGCGGGCGGCTGTGAATCTCAGCTGCTCAGAGAGCGGAGACACAAGAATTCATTTAA 660 |
| DB | 6600 ATGTTGGCGGGCGGCTGTGAATCTCAGCTGCTCAGAGAGCGGAGACACAAGAATTCATTTAA 6659 |
| QY | 661 ACCCAGAGAGGTGAGAGTTTGACAGTGCAGATGTGTCACCTCACTCCAACCTGGAGAC 720 |
| DB | 6660 ACCCAGAGAGGTGAGAGTTTGACAGTGCAGATGTGTCACCTCACTCCAACCTGGAGAC 6719 |
| QY | 721 AGAGTGAACATTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAACAGGTGAAG 780 |
| DB | 6720 AGAGTGAACATTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAACAGGTGAAG 6779 |
| QY | 781 TTAACCTTAAATACCCAAATGATCCCAATTCATCTTTCAAGTGTAAATTTAATATATAA 840 |
| DB | 6780 TTAACCTTAAATACCCAAATGATCCCAATTCATCTTTCAAGTGTAAATTTAATATATAA 6839 |
| QY | 841 ACAATTATGATGATGATACCTTATCATCTTTTCTTGTTCATATTAAGCTTTGAAGT 900 |
| DB | 6840 ACAATTATGATGATGATACCTTATCATCTTTTCTTGTTCATATTAAGCTTTGAAGT 6899 |
| QY | 901 GAGTATATATGTATATGCTGACAGACATCTCAATTGAGCTAGGTACATTTACAGTGTCTC 960 |
| DB | 6900 GAGTATATATGTATATGCTGACAGACATCTCAATTGAGCTAGGTACATTTACAGTGTCTC 6959 |

[illegible]

| | | | |
|-------------|---|-----------------------------------|-------|
| DB | 26147 | GCCAGTGGCTACTGTTGGATACATTAATTCAGA | 26112 |
| RESULT 8 | | | |
| ACAA60949/C | standard; DNA; 55827 BP. | | |
| ID | ACAA60949 | | |
| XX | ACAA60949; | | |
| AC | | | |
| XX | 11-AUG-2003 (first entry) | | |
| DT | | | |
| XX | DNA encoding human carboxypeptidase. | | |
| DE | | | |
| XX | | | |
| KW | Human; gene; carboxypeptidase; inflammation; cancer; an | | |
| KM | neurodegenerative disease; protease; ds; single nucleot | | |
| XX | SNP. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| PH | | | |
| Key | Location/Qualifiers | | |
| variation | replace(858,T) | | |
| FT | /*tag= u | | |
| FT | /note= "Single nucleotide polymorphism | | |
| FT | 2017. .53409 | | |
| CDS | /*tag= a | | |
| FT | /product= "Carboxypeptidase" | | |
| FT | 2017. .2090 | | |
| exon | /*tag= b | | |
| FT | /number= 1 | | |
| FT | 2091. .19206 | | |
| intron | /*tag= c | | |
| FT | /number= 1 | | |
| FT | replace(2122,T) | | |
| variation | /*tag= v | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(4088,A) | | |
| variation | /*tag= w | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(4240,A) | | |
| variation | /*tag= x | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(4416. .4418,TA) | | |
| variation | /*tag= y | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(4434,G) | | |
| variation | /*tag= z | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(4969,C) | | |
| variation | /*tag= aa | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(5868,A) | | |
| variation | /*tag= ab | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(6254,C) | | |
| variation | /*tag= ac | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(1017,C) | | |
| variation | /*tag= ad | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(10455,A) | | |
| variation | /*tag= ae | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(11613,A) | | |
| variation | /*tag= af | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(12130,T) | | |
| variation | /*tag= ag | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(17867. .17869,GA) | | |
| variation | /*tag= ah | | |
| FT | /note= "Single nucleotide polymorphism" | | |
| FT | replace(18243,G) | | |
| variation | /*tag= ai | | |
| FT | | | |

```
FT /note= "Single nucleotide polymorphism"
FT 19207. .19282
FT /*tag= d
FT /number= 2
FT 19283. .22683
FT /*tag= e
FT /number= 2
FT replace(22450,C)
FT /*tag= aj
FT /note= "Single nucleotide polymorphism"
FT 22684. .22808
FT /*tag= f
FT /number= 3
FT 22809. .24477
FT /*tag= g
FT /number= 3
FT replace(23003,T)
FT /*tag= ak
FT /note= "Single nucleotide polymorphism"
FT replace(24055. .24057,GA)
FT /*tag= al
FT /note= "Single nucleotide polymorphism"
FT replace(24132,T)
FT /*tag= am
FT /note= "Single nucleotide polymorphism"
FT 24478. .24586
FT /*tag= h
FT /number= 4
FT 24587. .26129
FT /*tag= i
FT /number= 4
FT replace(25713,A)
FT /*tag= an
FT /note= "Single nucleotide polymorphism"
FT 26130. .28231
FT /*tag= j
FT /number= 5
FT 28232. .33052
FT /*tag= k
FT /number= 5
FT replace(28476,C)
FT /*tag= ao
FT /note= "Single nucleotide polymorphism"
FT replace(29404,G)
FT /*tag= ap
FT /note= "Single nucleotide polymorphism"
FT replace(31047,T)
FT /*tag= aq
FT /note= "Single nucleotide polymorphism"
FT replace(31445,C)
FT /*tag= ar
FT /note= "Single nucleotide polymorphism"
FT replace(31447,C)
FT /*tag= as
FT /note= "Single nucleotide polymorphism"
FT replace(31600,C)
FT /*tag= at
FT /note= "Single nucleotide polymorphism"
FT replace(31714,C)
FT /*tag= au
FT /note= "Single nucleotide polymorphism"
FT replace(31715,T)
FT /*tag= av
FT /note= "Single nucleotide polymorphism"
FT replace(32193,G)
FT /*tag= aw
FT /note= "Single nucleotide polymorphism"
FT replace(32341,A)
FT /*tag= ay
FT /note= "Single nucleotide polymorphism"
FT replace(32341,G)
FT /*tag= ax
FT /note= "Single nucleotide polymorphism"
```

```
FT variation
FT replace(32561,G)
FT /*tag= az
FT /note= "Single nucleotide polymorphism"
FT replace(32600. .32602,AA)
FT /*tag= ba
FT /note= "Single nucleotide polymorphism"
FT replace(32642,T)
FT /*tag= bb
FT /note= "Single nucleotide polymorphism"
FT replace(32793. .32795,TG)
FT /*tag= bc
FT /note= "Single nucleotide polymorphism"
FT 33053. .33157
FT /*tag= 1
FT /number= 6
FT replace(33071,A)
FT /*tag= bd
FT /note= "Single nucleotide polymorphism"
FT 33158. .42288
FT /*tag= m
FT /number= 6
FT replace(34721,T)
FT /*tag= be
FT /note= "Single nucleotide polymorphism"
FT replace(35304,A)
FT /*tag= bf
FT /note= "Single nucleotide polymorphism"
FT replace(35425,C)
FT /*tag= bg
FT /note= "Single nucleotide polymorphism"
FT replace(36050,G)
FT /*tag= bh
FT /note= "Single nucleotide polymorphism"
FT replace(36291,G)
FT /*tag= bi
FT /note= "Single nucleotide polymorphism"
FT 42289. .42382
FT /*tag= n
FT /number= 7
FT 42383. .48643
FT /*tag= o
FT /number= 7
FT 48644. .48846
FT /*tag= p
FT /number= 8
FT 48847. .51176
FT /*tag= q
FT /number= 8
FT 51177. .51264
FT /*tag= r
FT /number= 9
FT 51265. .53224
FT /*tag= s
FT /number= 9
FT 53225. .53406
FT /*tag= t
FT /number= 10
FT XX
```

PN US2003017574-A1.

Query Match 12.0%; Score 240.2; DB 8; Length 55827;
Best Local Similarity 68.6%; Pred. No. 3.3e-44;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

```
QY 363 TCATACCTTAGAGTGTGTGTTTGTAGAAATATGCTGAGCTGCTTATGTCATTTCGA 422
   |||||
DB 46943 TGAATCCCTCAGACCGGGTGTATAGAAATATATGCAAGTCACATATTACTTTAA 46884
   |||||
QY 423 GTTTTGTAGCACCATTAAACAGTAAAC--AAGGCTGGCGCAGTGCTCACACT 480
   |||||
DB 46883 ATTTCAGTTTACCACTTTAAGAAAATATAAAGAGCGCAGCGTGCTTCACACT 46824
   |||||
QY 481 GTAATCCAGCACTTTGGAGGCTGAGCGAGCAGATCACTTTGGTCAGAGTTTGAGA 540
```


46823 ATATCCACACACTTTGGAGGCTGAGGAGCCGATCATCTTGATCAGAGTTTGACA 46764
46823 ATATCCACACACTTTGGAGGCTGAGGAGCCGATCATCTTGATCAGAGTTTGACA 46764
QY 541 CTAGCTGGCCCAACATGCGCAAACTGTGCTCTTAAAAAAATATCAAAATAGCCGTGC 600
46763 CCAGCGAGGCCAACATGCTGTAACCCCGCTCTACCAAAAA--TACAAAAATAGCCAGGC 46706
QY 601 ATGTGGCGGGCGCGCTTAATCTCAGCTGCTCAGAGAGCCGAGACAAAGATCATTAA 660
46705 ATAGTGTGACGCGCTTAATCCAGCTACTTGGAGGCTGAGGACAGAACTTGCTTGA 46646
QY 661 ACCGAGAGGTGAGGTGAGTTCAGTGAATGCTGCACT--CACTCAAACTGGAGAGA 719
46645 ACCGAGAGGACAGAGTGTGAGTGAATGCTGAGATCATGACACTGCACTCCAGCCGTGATGA 46586
QY 720 CAGAGTGACA-CTTTGTCTTCAAAAAAGAAAAAACAATGAAAAAGAACAGGTGA 778
46585 CAGAAACAGAACCCCTGCACACACACACACAAAGAAATATGTGAGATTAAATTCATA 46526
QY 779 AGTTAACTTAAATACCAATGTATCCCAAAATACATCATTTCAAAAGTGAATTAATA 838
46525 TTGTAAATTTAATTAACCAAAATGTATCAAAAGTATCA---TTTCAAGTGAACCAATATA 46470
QY 839 AAACAATTATGAATGAGATACCTTACATTTCTTTCTTTCTTTTCAATTAAGTCTTGAAA 898
46469 AA--AATTATCTTGAATAATTTATGTACTT-----TTTCACTACTAAGCTTCAAT 46417
QY 899 GTGATATATATGTTATGCTGACAGACATCTCAATTTGACATGACTAATTTCAGGTGC 958
46416 TCAGGTGTGATTAATTAACATCAGTGAATGTCTCAATTCAGACAGATCAATTTTAACAC 46357
QY 959 TCAGTACCCACATGTGCTGACACTTACTGTATTGGA 995
46356 TGGATAGCCACATGTGTTAGTGTGCTACTATGTGGA 46320
Db

RESULT 9
ABX13671/C
ID ABX13671 standard; DNA; 55827 BP.
XX
AC ABX13671;
XX
DT 14-FEB-2003 (first entry)
XX
DE Human protease gene.
XX
KM Human; gene; ds; protease; proteolytic degradation; proteolysis;
KM proliferation; differentiation; signaling; therapeutic; gene therapy;
KM protein therapy; diagnostic; immune response; vaccine; inflammation;
KM cancer; arteriosclerosis; degenerative disorder; chromosome 13;
KM single nucleotide polymorphism; SNP.
XX
XX Homo sapiens.
XX
FH Key location/Qualifiers
FT variation replace(858,T)
FT /*tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 2017..53409
FT /*tag= a
FT /product= "Protease"
FT exon 2017..2090
FT /*tag= b
FT /number= 1
FT intron 2091..19206
FT /*tag= c
FT /number= 1
FT variation /cons_splice= (5'site:yes,3'site:no)
FT replace(2122,T)
FT /*tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT replace(4088,A)
FT /*tag= w

FT variation /standard_name= "Single nucleotide polymorphism"
FT replace(4240,A)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ac
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ad
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ah
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT exon 19207..19282
FT /*tag= d
FT /number= 2
FT intron 19283..22683
FT /*tag= e
FT /number= 2
FT variation /cons_splice= (5'site:yes,3'site:no)
FT replace(22450,C)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism"
FT exon 22684..22808
FT /*tag= f
FT /number= 3
FT intron 22809..24477
FT /*tag= g
FT /number= 3
FT variation /cons_splice= (5'site:yes,3'site:no)
FT replace(23003,T)
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT variation /standard_name= "Single nucleotide polymorphism"
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT exon 24478..24586
FT /*tag= h
FT /number= 4
FT intron 24587..28129
FT /*tag= i
FT /number= 4
FT variation /cons_splice= (5'site:yes,3'site:no)
FT replace(25713,A)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism"

| | | | | |
|----|-----------|--|------------|--|
| FT | exon | 28130..28231 | /**tag= j | |
| FT | | /number= 5 | | |
| FT | intron | 28232..33052 | /**tag= k | |
| FT | | /number= 5 | | |
| FT | variation | replace(28476,C) | /**tag= ao | |
| FT | | /standard_name= replace(29404,G) | | |
| FT | variation | /**tag= ap | | |
| FT | | /standard_name= replace(3104,T) | | |
| FT | variation | replace(3104,T) | /**tag= aq | |
| FT | | /standard_name= replace(31445,C) | | |
| FT | variation | replace(31445,C) | /**tag= ar | |
| FT | | /standard_name= replace(31447,C) | | |
| FT | variation | replace(31447,C) | /**tag= as | |
| FT | | /standard_name= replace(31600,C) | | |
| FT | variation | replace(31600,C) | /**tag= at | |
| FT | | /standard_name= replace(31714,C) | | |
| FT | variation | replace(31714,C) | /**tag= au | |
| FT | | /standard_name= replace(31715,T) | | |
| FT | variation | replace(31715,T) | /**tag= av | |
| FT | | /standard_name= replace(32139,G) | | |
| FT | variation | replace(32139,G) | /**tag= aw | |
| FT | | /standard_name= replace(32341,R) | | |
| FT | variation | replace(32341,R) | /**tag= ax | |
| FT | | /standard_name= replace(32561,G) | | |
| FT | variation | replace(32561,G) | /**tag= ay | |
| FT | | /standard_name= replace(32600..35601,AAA) | | |
| FT | variation | replace(32600..35601,AAA) | /**tag= az | |
| FT | | /standard_name= replace(32642,T) | | |
| FT | variation | replace(32642,T) | /**tag= ba | |
| FT | | /standard_name= replace(32793..32794,TAG) | | |
| FT | variation | replace(32793..32794,TAG) | /**tag= bb | |
| FT | | /standard_name= replace(33053..33157) | | |
| FT | exon | 33053..33157 | /**tag= 1 | |
| FT | | /number= 6 | | |
| FT | variation | replace(33071,A) | /**tag= bc | |
| FT | | /standard_name= replace(33158..42288) | | |
| FT | intron | 33158..42288 | /**tag= m | |
| FT | | /cons_splice= 5'site:no,3'site:no) | | |
| FT | variation | replace(34721,T) | /**tag= bd | |
| FT | | /standard_name= replace(35304,A) | | |
| FT | variation | replace(35304,A) | /**tag= be | |
| FT | | /standard_name= replace(35425,C) | | |
| FT | variation | replace(35425,C) | /**tag= bf | |
| FT | | /standard_name= replace(36050,G) | | |
| FT | variation | replace(36050,G) | /**tag= bg | |
| FT | | /standard_name= replace(36291,G) | | |
| FT | variation | replace(36291,G) | /**tag= bn | |
| FT | | /standard_name= "Single nucleotide polymorphism" | | |

| | | | | |
|---|--------|--|------------------------|--|
| FT | exon | 42289..42382 | /+tag= n | |
| FT | | /number= 7 | | |
| FT | intron | 42383..48643 | /+tag= o | |
| FT | | /number= 7 | | |
| FT | exon | /coms_splice= 48644..48846 | (5'site:yes,3'site:no) | |
| FT | | /+tag= p | | |
| FT | | /number= 8 | | |
| FT | intron | 48847..51176 | /+tag= q | |
| FT | | /number= 8 | | |
| FT | exon | /coms_splice= 51177..51264 | (5'site:yes,3'site:no) | |
| FT | | /+tag= r | | |
| FT | | /number= 9 | | |
| FT | intron | 51265..53224 | /+tag= s | |
| FT | | | | |
| Query Match 12.0%; Score 240.2; DB 10; Length 55827; | | | | |
| Beat Local Similarity 68.6%; Pred. No. 3.3e-44; | | | | |
| Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7; | | | | |
| QY | | 363 TCATACCCCTGAGTAGTGGTGTGTTAGTAGAATAATAGCTGAGCTGTTATGTCATTCC | 422 | |
| DB | | 46943 TGAATCCTCGACCCAGGGTGTGTTAATAGAAATATATGCAAGTCACATATTTACTTTAA | 46884 | |
| QY | | 423 GTTTTATAGAGCCACATTAAACAGGTAA--MAGGCGGGCGGAGTGCTCACACT | 480 | |
| DB | | 46883 ATTTCAGTTACCACTTTAAGAAAATATAAGAGGCCAGGCACGGTGGCTCACACT | 46824 | |
| QY | | 481 GTAATCCAGCACTTTGGAGGCGTGAAGCAGGCAATCACCTTGGTCAGAGTTTGA | 540 | |
| DB | | 46823 ATATCCACAGACTTTGGAGGCGTGAAGCAGGCGCATCTTGAATCAGAGTTTGA | 46764 | |
| QY | | 541 CTAGCTGGCCACATGCGCAAACTGTGTCTTAAAAAAAATCAAAAATTAGCTCGC | 600 | |
| DB | | 46763 CCACCGAGGCCAACATGTGTAAACCCGCTCTCACCAAAA--TCAAAAATAGCCAGGC | 46706 | |
| QY | | 601 ATGGTGGCGGGCGCTGTATCTCAGCTGTCAGAGGCGCGAGACCAAGATCACTTAA | 660 | |
| DB | | 46705 ATAGTGGTGAACGCTGTATATCCAGCTACTTGGGAGGCTGAGGACGAACTTGCTTGA | 46646 | |
| QY | | 661 ACCGAGAGGTGAGGTTGCGATGAGCTGAGATCGTGCACCT--CACTCCAACTTGGAGA | 719 | |
| DB | | 46645 ACCGAGAGGCGAGGTTGCGATGAGCTGAGATCAGACACTCGCACTCCAGCTGGATGA | 46586 | |
| QY | | 720 CAGAGTGACA-CTTTGTCTCAAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA | 778 | |
| DB | | 46585 CAGAAACAAGACCTGCGACACACACACACAAAAAAGAAATATGTAGATTATTAATTA | 46526 | |
| QY | | 779 AGTTAACTTTAATACCCAAATGATCCCAATCAATCAATTTCAAGSTAAATTAATATA | 838 | |
| DB | | 46525 TTGTAAATTTATTAAACCAATATGTCAAAAGTATCA---TTTCAAGTAAACCAATATA | 46470 | |
| QY | | 839 AAACAATTAAGATAGATACCTTTAATCTTTTCTTGTTTTCATATTAAGCTTTGAAA | 898 | |
| DB | | 46469 AA--AATTATCTTGAATAATTTTAATGTACTT-----TTTTCATATCAAGCTTCAAAAT | 46417 | |
| QY | | 899 GTGAGTATATATGTTATGTGTGACAGACACATTCATATTGAGTACGTACATTTCAAGGTGC | 958 | |
| DB | | 46416 TCTGTGTGTAATTTAAACACTCAGATGATGTCCTCAATTTCAGACCAAGTACATTTTAAGAC | 46357 | |
| QY | | 959 TCAGTAGCCACATGTGGCTAGCACTGATCTGATTTGA | 995 | |
| DB | | 46356 TGGATAGCCACATGTGGTTAGTGGCTACTAATGCTGGA | 46320 | |
| Result 10 | | | | |
| ADS36454 | | | | |
| ID ADS36454 standard; DNA; 58337 BP. | | | | |
| XX | | | | |

| Query Match | 12.0%; Score 240.2; DB 13; Length 58337; |
|---|--|
| Best Local Similarity | 68.6%; Pred. No. 3.3e-44; |
| Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7; | |
| 363 TCATATCCCTGAGTAGTGTGTTTAAAGTAATATGCTGAGTCTTATGTCATTTC | 422 |
| Db 13209 TGAATCCTCAGACGAGGGTGTGTTTAAATAGAAATATATGCAATCATATTTACTTAA | 13268 |
| 423 GTTTTATAGTAGGCACATTTAAACAGGTAA--AAGCGTGGCGGAGTGGCTCACACT | 480 |
| Db 13269 ATTTTCAGATTACACTTTTAAAGAAATATTAAGAAAGAGCGGACATGGTGGCTCACACT | 13328 |
| 481 GTAATCCGAGCACTTTGGAGGAGCTGAGCAGCAGATCATCTTTGGTCAAGATTGGACA | 540 |
| Db 13329 ATATATCCAGCACTTTGGAGGAGCTGAGCAGGCGGATCATCTTGAATCAAGAGTTGGACA | 13388 |
| 541 CTAGCTGGCCAAATGCGCAACTCTGTCTCTTAAAAAAATACAAAATTAAGCTTGGC | 600 |

| | | | |
|-----------|--|--|-------|
| Db | 13389 | CCAGCGAGGCGCAACATGGTGAACCCCGCTCTCCACAAA--TACAAAATAAGCCAGGC | 13446 |
| Oy | 601 | ATGTGTGCGGGCGCCCTGTAACTCTCAGCTGCTCAGAGGCGCGAGACACAAAGATCACTTAA | 660 |
| Db | 13447 | ATAGTGTGTGACCCCTGTAACTCCAGCTACTTGGGAGGCTGAGGACGMAACTTGTCTTGA | 13506 |
| Oy | 661 | ACCCAGGAGGTGAGGTTGCGAGTGAGCTAGATCGGCGACT-CACTCCAACTGGAGGA | 719 |
| Db | 13507 | ACCCAGGAGGCGAGAGTTGCGAGTGAGCTAGATCATGACACTGCACTCCAGCTGGATGA | 13566 |
| Oy | 720 | CAGAGTGACA-CTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA | 778 |
| Db | 13567 | CAGAACAAAGACCTGCGACACACACACAAAAAGAAATGTGAGATTAAATTCATTA | 13626 |
| Oy | 779 | AGTTAACTTAAATPACCCAAATGATPCCAAATPACATCACTTCCAAAGTAAATTAATTA | 838 |
| Db | 13627 | TGTAAATTTATTTAAACCAATATGTACAAAGTATCA---TTTCAAGTGTAAACAAATTA | 13682 |
| Oy | 839 | AAACAATTAAGATGAGATPACTTACATCTTTTCTGTGTTTCATATTAAGCTTTGAA | 898 |
| Db | 13683 | AA-AAATTAATCTTGAAATATTTTAATGTTACTT---TTTCAATPCTAAGCCTTCAAAT | 13735 |
| Oy | 899 | GTGAGTATATATGTATTGCTGACAGACACATCTCAATTTGGACTAGCTACATTTCAAGTGC | 958 |
| Db | 13736 | TCGTGTGTGATTTTAACATCTCAGCTGATGTCTCAATTCAGACAGCTACATTTTAAGCAC | 13795 |
| Oy | 959 | TCAGTAGCCACATGTGGCTAGCAGTTACTGTATTGGA | 995 |
| Db | 13796 | TGATATGCCACATGTGGTTAGTGGCTATATGCTGGA | 13832 |
| RESULT 11 | | | |
| ID | AD536462/c | AD536462 standard; DNA; 64423 BP. | |
| XX | AD536462; | | |
| DT | 16-DEC-2004 | (first entry) | |
| DE | Human autoimmune disease-related genomic DNA sequence - SEQ ID 1676. | | |
| KW | single nucleotide polymorphism detection; SNP detection; | | |
| KW | rheumatoid arthritis; type 1 diabetes; multiple sclerosis; | | |
| KW | systemic lupus erythematosus; inflammatory bowel disease; psoriasis; | | |
| KW | thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo; | | |
| KW | glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease; | | |
| KW | primary systemic vasculitis; ds. | | |
| OS | Homo sapiens. | | |
| PN | MO2004083403-A2. | | |
| PD | 30-SEP-2004. | | |
| PF | 18-MAR-2004; 2004MO-US008461. | | |
| PR | 18-MAR-2003; 2003US-045544P. | | |
| PR | 25-APR-2003; 2003US-0465241P. | | |
| PA | (APPL-) APPLERA CORP. | | |
| PI | Cargill M, Begovich AB, Alexander HC; | | |
| DR | WPI; 2004-728480/71. | | |
| PT | New isolated nucleic acid molecule comprises at least 8 contiguous | | |
| PT | nucleotides where one of the nucleotides is a single nucleotide | | |
| PT | polymorphism (SNP), useful for diagnosing or treating autoimmune | | |
| PT | diseases, e.g. rheumatoid arthritis. | | |
| BS | Claim 16; SEQ ID NO 1676; 123bp; English. | | |
| CC | The invention comprises amino acid and coding sequences containing | | |

CC genetic polymorphisms associated with an altered risk of developing an
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further
CC comprises a method of identifying an individual that has an altered risk
CC of developing an autoimmune disease, comprising detecting a single
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA
CC and protein sequences of the invention are useful for diagnosing and
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease, the
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The
CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.

XX
XX
SQ Sequence 64423 BP; 19548 A; 12552 C; 12463 G; 19737 T; 0 U; 123 Other;

Query Match 12.0%; Score 240.2; DB 13; Length 64423;
Best Local Similarity 68.6%; Pred. No. 3.4e-44;

Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTAGAGTAGTGTGTTAGTAGAATAATAGCTAGCTGCTTATGATTCCTCA 422
DB 51306 TGAATCTCAGACAGGGTGTGTTAATAGAAATATATGCAAGTCACATATTTACTTTAA 51247
QY 423 GTTTTATAGAGCCACATTAATAACAGGTAA--AAGGCTGGGCGAGTGGCTCACACCT 480
DB 51246 ATTTTCAAGTTACACTTTTAAGAAAATAAAGAAAGGCGACGCGTGCCTCAACCT 51187
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGCGAGCGACATCCTTTGCTCAGAGCTTTGAGA 540
DB 51186 ATATCCCAACACTTTGGAGGCTGAGGCGAGCGAGTCACTTGAATCAGAGTGTGACA 51127
QY 541 CTAGCCTGGCAACATGCGAAACTGTCTCTAATAAATAAATAAATTTGCTGCTGCG 600
DB 51126 CAGGCGAGGCCAACATGCTAACCCTGCTTCAACAAAA--TACAAAAATGCGCAGGC 51069
QY 601 ATGCTGGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAAATCATTAA 660
DB 51068 ATAGTGTGACACCTGTATATCCAGTCACTTTGGAGGCTGAGGCGAGAACTTGCTTGA 51009
QY 661 ACCGAGAGGTGAGAGTGTGAGTGAAGTGTGCTCACT--CACTCCAACTGGGAGA 719
DB 51008 ACCGAGAGGTGAGAGTGTGAGTGAAGTGTGATCATGACACTCCAGCTGGATGA 50949
QY 720 CAGAGTGACA--CTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
DB 50948 CAGAAACAAGACCTGCGACACACACACAAAAAATATGTGATTAATTCATA 50889
QY 779 AGTTAATTTAATTAACCAATGTATCCCAATACAAATCATTTCMAAGTGAATTAATA 838
DB 50888 TTGTAATTTAATTAACCAATATGTCAAAATGCA----TTTCAAGTGAACCAATATA 50833
QY 839 AAAACAATTAATGAATGAGATACCTTCACTCTTTTCTGTTTCAATTAAGCTTTGAAA 898
DB 50832 AA--AATATCTTCTGAATATTTTATGTACTT-----TTTCAATTAAGCTTTCAAT 50780
QY 899 GTAGATATATATGTATGCTGAGACGACATCTCAATTTGAGTAGTACATTTCAAGTGC 958
DB 50779 TCTGGGTGTATTAACATCAAGTATGTCTCAATTCAGACAAAGTACATTTTAAGCAC 50720
QY 959 TCAGTAGCCACATGTGGCTAGCAGTTACTGTAATTTGA 995
DB 50719 TGGATAGCCACATGTGGTATGCTATGCTGGA 50683

RESULT 12

ACN45066/c
ID ACN45066 standard; DNA; 27189 BP.

XX ACN45066;

DT 18-NOV-2004 (first entry)

XX
XX
DE Human genomic sequence hCG30694.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

OS Homo sapiens.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1828; opp; English.

XX
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates, (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

XX
SQ Sequence 27189 BP; 6410 A; 6780 C; 6720 G; 7279 T; 0 U; 0 Other;

Query Match 11.8%; Score 235.6; DB 11; Length 27189;
Best Local Similarity 68.8%; Pred. No. 3e-43;

Matches 385; Conservative 0; Mismatches 164; Indels 11; Gaps 4;

QY 439 ATTAATAACAGTAAATAAGCTGGGCGAGTGGCTCACTGTATATCCAGCACTTTGG 498
DB 11830 AGTAAATATGCTAAATAATGGGCGGCGATGCTCATCTGTATATCCACACTATGA 11771
QY 499 GAGGCTGAGGCGAGGAGATCACTTTGTCAGAGATTGAGACTGAGCTGGGCAACATGG 558
DB 11770 GAGGCGAGGAGGCGAGATCACTTGAAGTCAAGAGTTTGAACAGCTGGGCAACATGG 11711
QY 559 CGAAACTGTCTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 618
DB 11710 TGAATCCCGTCTCT--ACTAAAGATATAAATAATAGTGGAGTGGTGGGTGCTCT 11652
QY 619 AATTCAGTGTCTCAGAGGCGGAGACACAGAAATCACTTAAACCCAGAGGTGAAGTT 678
DB 11651 AATCCAGTACTCAAGAGGCTGAGGAGAGAAATTTGCTGAACCCAGAGGAGAGGTT 11592
QY 679 GAGAGTGTGAGATCTGCGCACTCACTCCAACTGGGAGAGAGATGCACTTTTGTCT 738
DB 11591 GAGAGTGTGAGATCTGCGCACTCACTCCAACTGGGAGAGAGAGGAGAAATGCTCT 11532
QY 739 CAAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 798
DB 11531 CAAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 11477
QY 799 TGTATCCAAATACATCAATTTCAAGTGTAAATTAATTAATAAATAAATAAATAAATAAATA 858

| | | | |
|----|-------|---|--------|
| Db | 11476 | TACACCTCCCTCCCTTCATTTCAAAAAAATTTTCATGCAATTAAACAAACATTATTAA | 11417 |
| Qy | 859 | CTTTACACTTCCTTTCTTGTGTTTCATATTAAGCTTTGAAAGTAGATATATATGTC- | 917 |
| | | | |
| | | | |
| Db | 11416 | GATTACATCTC-----TTCCTCTCATGTATAGCTTTGAAATCTGGGGTATATTTATACT | 11361 |
| Qy | 918 | TGACAGACACTTCATTTTGGACTAGTAAATTTCAAGTGTCTCAGTACGCACATGTGGCT | 977 |
| Db | 11360 | TAAAGGATGCTCTTAAATTTGGACTAGACACATTTCAAGTGTCTCAGTACGCATGAT | 113010 |
| Qy | 978 | AGCAGTACTGTATTGGATG | 997 |
| | | | |
| Db | 11300 | AGTGCTCATCTGTATTAGATG | 11281 |

RESULT 13
ACM43998_3/c
Constitution (4 of 7) of ACM43998 from base 30001 (Human genomic sequence hCG1778483.)
MP Sequence split into 7 fragments LOCUS ACM43998 Accession ACM43998
Database Name
Ref

| Fragment Name | Begin | End |
|---------------|--------|--------|
| ACN43998-0 | 1 | 110000 |
| WB | | |
| ACN43998-1 | 100001 | 210000 |
| WB | | |
| ACN43998-2 | 200001 | 310000 |
| WB | | |
| ACN43998-3 | 300001 | 410000 |
| WB | | |
| ACN43998-4 | 400001 | 510000 |
| WB | | |
| ACN43998-5 | 500001 | 610000 |
| WB | | |
| ACN43998-6 | 600001 | 653122 |

| | | | | |
|---------------------------|--------|--------------------|------------|----------------|
| Query Match | 11.5% | Score 229.2; | DB 11; | length 110000; |
| Best Local Similarity | 69.7%; | Pred. No. 1.3e-41; | | |
| Matches 404; Conservative | 0; | Mismatches 153; | Indels 23; | Gaps 6; |

| | | | |
|----|-------|---|-------|
| QY | 440 | TTAAACAGGTTAAAAAGGCTGGGGCCGAGTGGCTCAACCTGTAAATCCAGCACTTTGGG | 499 |
| Db | 60810 | TTTTAAAAAATACAAATTGGGCCAGAGTGGGGCTCATGTCTGTAAATCCAGCACTTTGAA | 60751 |
| QY | 500 | AGGCTGAGGAGGAGAGATCACTTTGGTCAAGATTGGAGACTAGGCTGGCCAAATGCG | 559 |
| Db | 60750 | AGGCGGAGGGGGGAAATCAC--GAGGTCAAGATTGGAGCCAGGCTGACCAATGTG | 60693 |
| QY | 560 | GAACCTGTGTCTTAAAAAAAATCAAAAATTAGCTGGCATGTGGCGGGCCTGTGA | 619 |
| Db | 60692 | GAATCCCCGCTCT-ACTAAAAATTACAAAAATTAGCCAGGCGTGGTGGTCAAGCTGTGA | 60634 |
| QY | 620 | ATCTCAGCTCTCAGAGGCGCCGAGACACAAGATTCATTAAACCGAGAGGTGGAGCTGG | 679 |
| Db | 60633 | ATCCCAAGCCACTTGGGAGGCTGGAGACAGGAATCACTTGATCTTGGGAGGCGAGAAGTGG | 60574 |
| QY | 680 | CAGTGAAGCTAGATTCGTGCACCT-CACTCCAACTGGGAGACAGATGACACTTTGTCT | 738 |
| Db | 60573 | CAATGAATGAGATTATGCACACTGCACCTCCAGCTGGGCGACAGACAAAGATCCATCTG | 60514 |
| QY | 739 | CAAAAAGAAAAAAAACAGTAAAAAAGAAAACAGTGAGTTAACTTAAATAACCA | 798 |
| Db | 60513 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAATGG-----ATTTAACTTAA | 60466 |
| QY | 799 | TGATCCCAATCAATCATTTTCAAAAGTAAATTAATTAATTAACATTAATGAATGATA | 858 |
| Db | 60465 | TATATCCAAATATGTTTAAATTTCAACATGCAACTAATATTAAGAAATGAGATGTTA | 60406 |
| QY | 859 | CTTTCACATCTTTTCTGTTTCAATTAATTAAGCTCTTGAAGTAGATATATGTTAGCT | 918 |
| Db | 60405 | ACCTTCTTCTTCTTTTTC---TACTAAGTCTTTGGATCCAGCATATATTTTATACA | 60349 |
| QY | 919 | GACAGCACATCTCAATTGGACTAGCT---ACATTTCAAGTGTCTCAATGACCAATGTG | 974 |
| Db | 60348 | TATGGCTATCTGATATGGGGCTAGCTAGCCACGTTTCCAGGGCTCAAGAGCCACATCTG | 60289 |
| QY | 975 | GCTAGCAGTTACTGTATGTGATGGACCGGATCTTAAAGGA | 1014 |
| Db | 60288 | GCTGTGAGCTAACATACAGACAGTGCAGGACTTGAAGCA | 60249 |

| | |
|------------|---|
| RESULT 14 | |
| AAK68359/c | |
| ID | AAK68359 standard; DNA; 2953 BP. |
| XX | |
| XX | |
| AC | AAK68359; |
| XX | |
| DT | 06-NOV-2001 (first entry) |
| XX | |
| DE | Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23171. |
| XX | |
| KW | Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; |
| KM | cytostatic; gene therapy; vaccine; metastasis; da. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200157182-A2. |
| XX | |
| PD | 09-AUG-2001. |
| XX | |
| PF | 17-JAN-2001; 2001WO-US0001354. |
| XX | |
| PR | 31-JAN-2000; 2000US-0179065P. |
| PR | 04-FEB-2000; 2000US-0180628P. |
| PR | 24-FEB-2000; 2000US-0184664P. |
| PR | 02-MAR-2000; 2000US-0186350P. |
| PR | 16-MAR-2000; 2000US-0189874P. |
| PR | 17-MAR-2000; 2000US-0190076P. |
| PR | 18-APR-2000; 2000US-0198123P. |
| PR | 19-MAY-2000; 2000US-0205515P. |
| PR | 07-JUN-2000; 2000US-0209467P. |
| PR | 28-JUN-2000; 2000US-0214886P. |
| PR | 30-JUN-2000; 2000US-0215135P. |
| PR | 07-JUL-2000; 2000US-0216647P. |
| PR | 07-JUL-2000; 2000US-0216880P. |
| PR | 11-JUL-2000; 2000US-0217487P. |
| PR | 11-JUL-2000; 2000US-0217496P. |
| PR | 14-JUL-2000; 2000US-0218280P. |
| PR | 26-JUL-2000; 2000US-0220963P. |
| PR | 26-JUL-2000; 2000US-0220964P. |
| PR | 14-AUG-2000; 2000US-0224518P. |
| PR | 14-AUG-2000; 2000US-0224519P. |
| PR | 14-AUG-2000; 2000US-0225213P. |
| PR | 14-AUG-2000; 2000US-0225214P. |
| PR | 14-AUG-2000; 2000US-0225266P. |
| PR | 14-AUG-2000; 2000US-0225267P. |
| PR | 14-AUG-2000; 2000US-0225268P. |
| PR | 14-AUG-2000; 2000US-0225270P. |
| PR | 14-AUG-2000; 2000US-0225477P. |
| PR | 14-AUG-2000; 2000US-0225757P. |
| PR | 14-AUG-2000; 2000US-0225758P. |
| PR | 14-AUG-2000; 2000US-0225759P. |
| PR | 18-AUG-2000; 2000US-0226279P. |
| PR | 22-AUG-2000; 2000US-0226681P. |
| PR | 22-AUG-2000; 2000US-0226688P. |
| PR | 22-AUG-2000; 2000US-0227182P. |
| PR | 23-AUG-2000; 2000US-0227009P. |
| PR | 30-AUG-2000; 2000US-0228934P. |
| PR | 01-SEP-2000; 2000US-0229287P. |
| PR | 01-SEP-2000; 2000US-0229343P. |
| PR | 01-SEP-2000; 2000US-0229344P. |
| PR | 01-SEP-2000; 2000US-0229345P. |
| PR | 05-SEP-2000; 2000US-0229509P. |
| PR | 05-SEP-2000; 2000US-0229513P. |
| PR | 06-SEP-2000; 2000US-0230437P. |
| PR | 06-SEP-2000; 2000US-0230438P. |
| PR | 08-SEP-2000; 2000US-0231242P. |
| PR | 08-SEP-2000; 2000US-0231243P. |
| PR | 08-SEP-2000; 2000US-0231244P. |
| PR | 08-SEP-2000; 2000US-0231413P. |
| PR | 08-SEP-2000; 2000US-0231414P. |
| PR | 08-SEP-2000; 2000US-0232080P. |
| PR | 08-SEP-2000; 2000US-0232081P. |
| PR | 12-SEP-2000; 2000US-0231968P. |

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0234989P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241866P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250316P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 23171; 3071bp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;

Query Match 11.4%; Score 227.6; DB 4; Length 2953;
Best Local Similarity 69.0%; Pred. No. 1,1e-41;
Matches 421; Conservative 0; Mismatches 164; Indels 25; Gaps 7;

QY 414 TCATTTCCAGTTTATTTAGTACCCACATTAAACAGGTAAACAA-----GGCTGGCGCA 467
DB 1699 TAATTAAAGATTTTAAACAGCCACATTAATGAAGAAACAAAGTCGGCGGCGCG 1640
QY 468 GTGGCTCAACCTGTAAATCCAGCACTTTGGAGGTGAGGAGGAGATCACTTTGGT 527
DB 1639 GTGGCTCAACCTGTAAATCCAGCACTTTGGAGGTGAGGAGGAGATCACTTTGGT 1582
QY 528 CAGGAGTTTGAAGCTAGCTGTGGCAACATGGGAAACTGTCTCTCAAAAAAATATCAAA 587
DB 1581 CGGAAGATGAAACCATCTTGTAAACAGCGTAACTCCGTCTCAAAAAAATATCAAA 1522
QY 588 AAATTAGCTGGCATGTGGCGGCGCTGTAACTCACTCTCAGAGGCGGAGACAC 647
DB 1521 AAATTAGCGAGCATGTGGCGGAGCTGTAGTCCCACTACTTGGAGGCTGAGGACAG 1462
QY 648 AAGAACTCACTTAAACCCAGAGAGTGAAGTTGCACTGAGCTAGATCGGCCACTCAGTC 707
DB 1461 GAGAAATGGGTAAACCCGAGAGAGCTTCACTGAGTCACTGATGAGACCACTGCACT 1402
QY 708 CAACCTGGAGAGCAAGAGTACACTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAA 767
DB 1401 CACCAAGCTGGGCGAGAGAGC-AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAAA 1343
QY 768 GAAACAGGTGAAGTTAATTAATTAATCAACCAATGTATCCAAATACATCTTCAAAAGTG 827

Db 1342 TAAGTCATAGATTTAATATATTAGCTAGATATGCTAAATATATCATTTAA----- 1288
Qy 828 TAATTAATATAAACAATTATGATGATATCTTACATTTCTTTCTGTTTATATTA 887
Db 1287 -CATGTATACAAAGATTATTAACAATAATTTACATTCCTGTGG---GGGTATTA 1232
Qy 888 AGCTTTGAAAGAGATATATA-TGTTATGCTGACAGACATCTCAATTTGGAAGTACTA 946
Db 1231 AATCTTTGAATTAATTAATGTATTTTATTTACTTAAGCTGATCTTGATTCAGACTAGCCA 1172
Qy 947 CATTTCAGTGCTCAGTAGCCATGTGGCTAGCA-----GTTACTGTATTTGATGCA 1000
Db 1171 CATTTCAGTGCTCAGTAGCCCATGTGGCTAGCAAGATGATGATTAATGATGCA 1112
Qy 1001 CGGATCTTGA 1010
Db 1111 CAGCTTATTA 1102

RESULT 15
AAK68360/C
ID AAK68360 standard; DNA; 2953 BP.
XX AAK68360;
AC
XX
DT 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23172.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225447P.
XX
XX 14-AUG-2000; 2000US-0225757P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 16-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226681P.
XX
XX 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.

17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0254678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 23172; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I), proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 CC
 XX
 XX Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;

Query Match 11.4%; Score 227.6; DB 4; Length 2953;
 Best Local Similarity 69.0%; Pred. No. 1.1e-41;
 Matches 421; Conservative 0; Mismatches 164; Indels 25; Gaps 7;

QY 414 TCATTTCAGTTTATAGTACCATTAACAGGTAAAAA-----GGCTGGCGCA 467
 Db 1699 TAATTAAAGATTTTAAACGACATTAATTAATGTAAGAAAAAGTCGGCGCGCG 1640
 QY 468 GTGGCTCACACCTGTATATCCAGCACTTTGGAGGCTGAGGCAAGGCAATCCTTTGGT 527
 Db 1639 GTGGCTCACACCTGTATATCCAGCACTTTGGAGGCTGAGGCAAGGCAATCCTTTGGT 1582

QY 528 CAGGATTTGAGACTAGCTGGCAACATGGGAACTGTCTCTTAATAAATAATACAA 587
 Db 1581 CCGAAGATGAGACCATCTTGTAAACGCGGAACTCCGCTCTACTAATAATACAA 1522
 QY 588 AAATTAGCTGGCATGTGGCGGGCGCTGTAACTCACTGCTCAGAGGCGGAGACAC 647
 Db 1521 AAATTAGCCAGGCATGTGGCGGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGAG 1462
 QY 648 AAGAATCACTTAAACCCAGAGGTGAGGTTGCACTGAGCTGAGATCGGCCACTCATC 707
 Db 1461 GGAATGGGCTGAACCCGAGGAGGAGCTGAGGAGCTGAGATGAGCACTGCACT 1402
 QY 708 CAACCTGGAGACAGAGTGACACTTTTGTCAAAAAAGAAAAAACAAGTAAAAA 767
 Db 1401 CACCCAGCTGGGCGGACAGAGAC-AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAAA 1343
 QY 768 GAAACAGGTGAAGTTAATCTTAATAACCAATGTATCCAAATACATCTTCAAAAGTG 827
 Db 1342 TAAGTCATGATTTTAATATATTTAGCTCAGATGTCTAAATATATCATTTAA----- 1288
 QY 828 TAATTATATTAACAATTATGAGATTAATTACATCTTTCTTGTTCATATTA 887
 Db 1287 -CATGTATCAACAAAGTTATTAACAAATATTTCAATTCCTTGTGG--GGTAAATA 1232
 QY 888 AGTCTTGAAGAGTATATA-TGTTATGCTGACAGACATCTCAATTGGAAGTAACTA 946
 Db 1231 AATCTTGAATTAATATGTATTTTATTAATCAAGCTGATCTTGATTAAGCTAAGCA 1172
 QY 947 CATTCAGGTGCTCAGTACCAATGTGGCTAGCA-----GTACGTATTTGATGGCA 1000
 Db 1171 CATTCAGGTGCTCAGTACCAATGTGGCTAGCAATGTGTTACTATAGTAATAGCA 1112
 QY 1001 CGATCTTACA 1010
 Db 1111 CAGCTCTATA 1102

Search completed: February 25, 2005, 20:15:00
 Job time : 1024.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:28:50 / Search time 4913 Seconds

(without alignments)
2889.757 Million cell updates/sec

Title: US-09-936-271b-14

Perfect score: 1608
Sequence: 1 MATARPPMMVLCALITALL.....VYTNCKFTKIQETIQANS 293

Scoring table:

BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cg211/USPTO.spool/US09936271/runat.23022005.130127.27262/app_query.fasta_1.455
-DB=GenEmbl -QFMT=faetap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=spct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09936271.QCGN.1.1.3731.@runat.23022005.130127.27262 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1608 | 100.0 | 882 | 9 | BT006867 Homo sapi |
| 2 | 1608 | 100.0 | 882 | 12 | BT007831 Synthetic |
| 3 | 1608 | 100.0 | 1370 | 9 | AY279380 Homo sapi |
| 4 | 1608 | 100.0 | 1381 | 6 | BD107879 |
| | | | | | BD107879 36 human |

| | | | | | | |
|----|-------|-------|--------|----|----------|---------------------|
| 5 | 1608 | 100.0 | 1387 | 6 | CQ874884 | CQ874884 Sequence |
| 6 | 1608 | 100.0 | 1387 | 6 | CQ874960 | CQ874960 Sequence |
| 7 | 1608 | 100.0 | 1387 | 9 | AF168768 | AF168768 Homo sapi |
| 8 | 1608 | 100.0 | 1438 | 9 | AY279381 | AY279381 Homo sapi |
| 9 | 1608 | 100.0 | 1499 | 6 | AR352504 | AR352504 Sequence |
| 10 | 1608 | 100.0 | 1499 | 6 | BD005362 | BD005362 Protease |
| 11 | 1608 | 100.0 | 1516 | 6 | BD107865 | BD107865 36 human |
| 12 | 1608 | 100.0 | 1527 | 9 | BC008036 | BC008036 Homo sapi |
| 13 | 1608 | 100.0 | 1570 | 6 | AR252580 | AR252580 Sequence |
| 14 | 1608 | 100.0 | 1570 | 6 | AR528679 | AR528679 Sequence |
| 15 | 1608 | 100.0 | 1570 | 6 | AX080829 | AX080829 Sequence |
| 16 | 1608 | 100.0 | 1570 | 6 | AX464322 | AX464322 Sequence |
| 17 | 1608 | 100.0 | 1570 | 6 | AX464322 | AX464322 Sequence |
| 18 | 1608 | 100.0 | 1570 | 6 | AY359010 | AY359010 Homo sapi |
| 19 | 1595 | 99.2 | 1476 | 6 | AR078184 | AR078184 Sequence |
| 20 | 1595 | 99.2 | 1476 | 6 | AR137506 | AR137506 Sequence |
| 21 | 1595 | 99.2 | 1476 | 6 | AR242358 | AR242358 Sequence |
| 22 | 1595 | 99.2 | 1476 | 6 | BD082659 | BD082659 Keratinoc |
| 23 | 1571 | 97.7 | 1301 | 9 | AY461805 | AY461805 Homo sapi |
| 24 | 1513 | 94.1 | 1439 | 6 | BD107880 | BD107880 36 human |
| 25 | 1479 | 92.0 | 1143 | 6 | CQ720600 | CQ720600 Sequence |
| 26 | 1475 | 91.7 | 929 | 9 | AF435981 | AF435981 Homo sapi |
| 27 | 1475 | 91.7 | 1132 | 9 | AF435980 | AF435980 Homo sapi |
| 28 | 1403 | 87.3 | 1504 | 6 | AR263823 | AR263823 Sequence |
| 29 | 1002 | 62.3 | 11570 | 6 | CQ788219 | CQ788219 Sequence |
| 30 | 1002 | 62.3 | 11570 | 6 | CQ874885 | CQ874885 Sequence |
| 31 | 1002 | 62.3 | 11570 | 6 | CQ874961 | CQ874961 Sequence |
| 32 | 1002 | 62.3 | 11570 | 6 | AF135028 | AF135028 Homo sapi |
| 33 | 1002 | 62.3 | 217346 | 2 | AC027602 | AC027602 Homo sapi |
| 34 | 1002 | 62.3 | 230000 | 9 | AF243527 | AF243527 Homo sapi |
| 35 | 997 | 62.0 | 107487 | 2 | AC011483 | AC011483 Homo sapi |
| 36 | 985.5 | 61.3 | 200792 | 2 | AC130782 | AC130782 Pan trogl |
| 37 | 725 | 45.1 | 735 | 6 | AX429955 | AX429955 Sequence |
| 38 | 725 | 45.1 | 738 | 6 | BD139877 | BD139877 Compounds |
| 39 | 690.5 | 42.9 | 116508 | 2 | AC148431 | AC148431 Canis fam |
| 40 | 688 | 42.8 | 1140 | 6 | AX067353 | AX067353 Sequence |
| 41 | 676 | 42.0 | 152577 | 2 | ASU67256 | U76256 Sus scrofa |
| 42 | 675.5 | 42.0 | 152577 | 2 | AC148863 | AC148863 Sus scrofa |
| 43 | 675.5 | 42.0 | 215342 | 4 | AC149292 | AC149292 Sus scrofa |
| 44 | 674 | 41.9 | 1078 | 10 | BC063763 | BC063763 Mus muscu |
| 45 | 671 | 41.7 | 765 | 6 | CQ733642 | CQ733642 Sequence |

ALIGNMENTS

RESULT 1
BT006867
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BT006867 Homo sapiens kallikrein 5 mRNA, complete cds.
BT006867.1 GI:30582572
FLI CDNA.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 882)
Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M. and Farmer,A.
Cloning of human full-length cDNAs in BD Creator(TM) System Donor vector
Unpublished
2 (bases 1 to 882)
Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M. and Farmer,A.
Direct Submission
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA
This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two

forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

```

1..882
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="GH00284X1.0"
    /clone_1lb="BD Creator (TM) CDS Library derived from MGC collection"
    /lab_host="DH5alpha TI resistant"
    /note="Vector: pDNR-Dual"
1..882
    /codon_start=1
    /product="Kallikrein 5"
    /protein_id="AAP35513.1"
    /db_xref="GI:30582573"
    /translation="MATARPMMWVLCALITALLAVTHVLANNDVSCDHSNTVPS
    GSNQD/GAAGEDPARSDSSRLINSGDCDMHTQPMQALILRPNOLYCAVLVHPQW
    LITAAICRKRVFRLVGLHYSLSPVSGQMFQGVKSIPIHPGVSHGNSDMLIKLN
    RRIPTKDVAPINVSHPGSAIGKCLVSGMGTKSPQVHPKVLQCLNTSVLSORKE
    DAVPRQIDDTMPCAGDKAGRDSQGGSGGVVGNGLQGLVNSGDYPCARPRRPGVYT
    NICKTKMIOETIQANS"

```

CDS

```

1..882
    /product="Kallikrein 5"
    /protein_id="AAP35513.1"
    /db_xref="GI:30582573"
    /translation="MATARPMMWVLCALITALLAVTHVLANNDVSCDHSNTVPS
    GSNQD/GAAGEDPARSDSSRLINSGDCDMHTQPMQALILRPNOLYCAVLVHPQW
    LITAAICRKRVFRLVGLHYSLSPVSGQMFQGVKSIPIHPGVSHGNSDMLIKLN
    RRIPTKDVAPINVSHPGSAIGKCLVSGMGTKSPQVHPKVLQCLNTSVLSORKE
    DAVPRQIDDTMPCAGDKAGRDSQGGSGGVVGNGLQGLVNSGDYPCARPRRPGVYT
    NICKTKMIOETIQANS"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      7..98e-121      Length:      882
Score:          1608.00        Matches:      293
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%        Indels:        0
DB:              9             Gaps:            0

```

US-09-936-271B-14 (1-293) x BT006867 (1-882)

```

QY      1 MetAlaThrAlaArgProPOTPTpMetTTPValLeuCYaAlaLeuIleThrAlaLeuLeu 20
DB      1 ATGGCTACAGCAAGACCCCTGATGTGGTCTCTGCTCTGATCACAGCCTTGCTT 60
QY      21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
DB      61 CTGGGGGTACAGAGCATGTTCTGCCCAACATGATGTTCTCTGTACCACTCTTAC 120
QY      41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGluAspAlaArgSer 60
DB      121 ACCGTGCTCTGTGGAGCAACCGAGCTGGAGCTGGGGCCGGGGAGACCGCCGGTGG 180
QY      61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTP 80
DB      181 GATGACAGACGACGCGCATCAATGATGATTCGACTGCGATGTGACACCCCGCGTGG 240
QY      81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValValLeuValHisPro 100
DB      241 CAGGCGCGCTGTGCTTACAGGCCCAACACACTTACTGCGGGCGGGTGTGGTCAATCCA 300
QY      101 GlnTPRLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
DB      301 CAGTGGCTGCTCAACGCGCCGCTGCAAGAGAAAGATTTCAGAGTCCGCTCTCGGCCAC 360
QY      121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
DB      361 TACTCCCTGTCAACCAATTATGATCTGGGCAACAATGTTTCCAGGGGGTCAAAATCATC 420
QY      141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValysAsn 160
DB      421 CCCCACTCTGGCTACCTCCACCTGGCACTCTTACAGACCTCAAGCTCAACAACTGAAAC 480
QY      161 ArgArgIleArgProThrIysAspValArgProIleAsnValSerSerHisCysProSer 180

```

```

DB      481 AGAAGATTGTCCTCCACTTAAGATGTCAGACCCATCAAGCTCTCTCATGTCCCTCT 540
QY      181 AlaGlyThrIlyCysLeuValSerGlyTTPGlyThrThrIlySerProGlnValHisPhe 200
DB      541 GCTGGGACAAATGCTTGTGTCTGTGGCTGGGGGACAAACCAAGACCCCAAGTGCATTC 600
QY      201 ProIysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlySerGlyIuAspAla 220
DB      601 CTTAAGTCTCTCAAGTCTTGAATATACCGCTGCTTACAGAAAAGTGCAGAGATGCT 660
QY      221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyAspIlyArgAspSer 240
DB      661 TACCCGACAGATGATGATACACCATGTTCTGCGCGGGGACAAAGACAGTACAGCTCC 720
QY      241 CysGlnGlyAspSerGlyIlyProValValCysAsnGlySerLeuGlnIlyLeuValSer 260
DB      721 TGCAGAGGTGATTCGGGGGGGCTGTGTGCTCAATGCTCCTCGACGGAGCTCGTTC 780
QY      261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlys 280
DB      781 TGGGAGATTAACCTTGTCTGCGGCCCAACAGACCGGGTGTCTTACAGAACTCTGCAAG 840
QY      281 PheThrIlySerIleGlnIlyThrIleGlnAlaAsnSer 293
DB      841 TTCACCAATGATTCAGAGAAACCATCCAGGCCAATCC 879

```

RESULT 2

BT007831

LOCUS BT007831 882 bp mRNA linear SYN 13-MAY-2003

DEFINITION Synthetic construct Homo sapiens kallikrein 5 mRNA, partial cds.

ACCESSION BT007831.1 GI:30584500

KEYWORDS FLI CDNA.

SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.

REFERENCE 1 (bases 1 to 882)

AUTHORS

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Pheilan,M. and Farmer,A.

TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor

JOURNAL

Unpublished

REFERENCE 2 (bases 1 to 882)

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Pheilan,M. and Farmer,A.

TITLE Direct Submision

JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow

Circle, Palo Alto, CA 94303, USA

COMMENT

This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

```

1..882
    /organism="synthetic construct"
    /mol_type="mRNA"
    /db_xref="taxon:32630"
    /clone="GH00284X1.0"
    /clone_1lb="BD Creator (TM) CDS Library derived from MGC collection"
    /lab_host="DH5alpha TI resistant"
    /note="Vector: pDNR-Dual"
1..882
    /note="Mutations: 881:Stop->Leu"

```

CDS

```

/codon_start=1
/translate="Homo sapiens kallikrein 5"
/protein_id="AAP36503.1"
/db_xref="GI:30584501"
/translation="MATARPMMVLCALITALLGVTEHVALNDVSCDPSNTVPS
GNSODIAGAGEDBARSDSSSRITNGSDCMHTOPQALLLPNOLYCAVALVHPQW
LTLAAHCRKRVPRVRLGHVSLSPVSGQMGQVKSIPHPGSHBGHNDMLIKLN
RRIKPTKVRPIVNSHCPASAGTKCLVSGMTTSPQVHPKVLQCLNLSVLSQKCE
DAYPRQIDDTMPCAGDKGRDSCQGSQSGPVVNCNGLQGLVSWGDYPCARPNNRGVYT
NICKFTKWIQETIOANSI"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7,386-121 Length: 882
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

US-09-936-271b-14 (1-293) x BT007831 (1-882)

```

QY 1 MetAlaThrAlaArgProTyrMetTyrValLeuCysAlaLeu1LeuThrAlaLeu 20
DB 1 ATGGCTACGACGACGACCCCTGGATGGGCTCTGTCTCTGATCAGCCTTCTT 60
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHis 40
DB 61 CTGGGGGCTCAGACGATGTTCTGCGCAACATGATTTCTGTGACCACTTCTAAC 120
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValAlaGlyValAspAlaArgSer 60
DB 121 ACCGTGCTCTGGAGACCAACGACCTGGAGCTGGGCGGGAAGCGCCGGTGG 180
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80
DB 181 GATGACGACGACGACCGGCATCATCAATGATCCAGTCCGATATGACACCGCCTGG 240
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValAlaLeuValHisPro 100
DB 241 CAGGCGCGCTGTGTGCTAAAGCCCAACCACTTACTGCGGGGGGTGTGGTGATCCA 300
QY 101 GlnTyrLeuLeuThrAlaAlaHisCysArgGlyValValIleAsnArgValHis 120
DB 301 CAGTGGCTGCTCAGCGCGCCCACTGACGAGAAAGATTTCAGAGTCCGTTCTCGGCAC 360
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIle 140
DB 361 TACTCCCTGTCAACAGTTTATGAAATCTGGGACAGATGTTCCAGGGGGTCAAAATCCATC 420
QY 141 ProHisProGlyTyrSerHisAspProGlyHisSerAsnAspLeuMetLeuIle 160
DB 421 CCCCACTCTGGCTACTCCACCTCGGCACCTTAACGACCTCACTCATCAACTGAAC 480
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 481 AGAAGAAATTCCTCCCACTAAAGATGTCAGACCACTCAACCTCTCTCAATGCTCTCT 540
QY 181 AlaGlyThrIleCysLeuValSerGlyTyrGlyThrThrIleSerSerProGlnValHisPhe 200
DB 541 GCTGGGACAAAGTCTTGGTGTCTGCTGGGACAAACCAAGACCCCAAGTGCATTC 600
QY 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
DB 601 CTTAAAGTCTCTCAAGTCTTGAATATCAGCTGTCTAAGTCAAAAAGTGCAGAGATGT 660
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaIleValAspValAlaGlyAspSer 240
DB 661 TACCCGAGACAGATGATGACACCATGTCTGCGCCGCTGACAAAGAGATGAGACTCC 720
QY 241 CysGlnGlyAspSerGlyIleProValIleCysAsnGlySerLeuGlnGlyLeuValSer 260

```

```

DB 721 TGGCAGGATGATTCCTGGGGGGCTGTGTGTGCAATGCTCCCTGACGAGGACTGTGTCT 780
QY 261 TTTGlyAspTyrProCysAlaAlaArgProAsnArgProGlyValIleThrAsnLeuCysVal 280
DB 781 TGGGAGATTAACCTTGTGCTGGCCGCCCAACGACCGGGTCTTCAACGAACTCTGCAAG 840
QY 281 PheThrIleTyrIleGlnGlnIleGlnIleGlnAlaAsnSer 293
DB 841 TTTCAACATGTGATCCAGAAACCATTCAGGCCACTCC 879

```

```

RESULT 3
LOCUS AY279380 1370 bp mRNA linear PRI 26-MAY-2003
DEFINITION Homo sapiens kallikrein 5 splice variant 1 (KUK5) mRNA, complete
cbs; alternatively spliced.
ACCESSION AY279380
VERSION AY279380.1 GI:31075480
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1370)
Kurlender, L., Yousef, G.M., White, N.M.A., Robb, J.-D., Borgono, C.A.
and Diamandis, E.P.
Identification of splice variants for the human kallikrein gene 5
(KUK5)
Unpublished
2 (bases 1 to 1370)
Kurlender, L., Yousef, G.M., White, N.M.A., Robb, J.-D., Borgono, C.A.
and Diamandis, E.P.
Submitted (19-Apr-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada

```

```

JOURNAL
REFERENCE
AUTHORS
TITILE
JOURNAL
FEATURES
source
Location/Qualifiers
1. .1370
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. .1370
/gene="KUK5"
1. .127
/gene="KUK5"
128. .1009
/gene="KUK5"
/note="serine protease; hKS; alternatively spliced;
similar to the product of GenBank Accession Number
AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 1"
/protein_id="AAP42275.1"
/db_xref="GI:31075481"
/translation="MATARPMMVLCALITALLGVTEHVALNDVSCDPSNTVPS
GNSODIAGAGEDBARSDSSSRITNGSDCMHTOPQALLLPNOLYCAVALVHPQW
LTLAAHCRKRVPRVRLGHVSLSPVSGQMGQVKSIPHPGSHBGHNDMLIKLN
RRIKPTKVRPIVNSHCPASAGTKCLVSGMTTSPQVHPKVLQCLNLSVLSQKCE
DAYPRQIDDTMPCAGDKGRDSCQGSQSGPVVNCNGLQGLVSWGDYPCARPNNRGVYT
NICKFTKWIQETIOANSI"
1010. .1370
/gene="KUK5"
1319. .1324
/gene="KUK5"

```

```

ORIGIN
polya_signal
Alignment Scores:
Pred. No.: 1,286-120 Length: 1370
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

DB: 9 Gaps: 0
US-09-936-271b-14 (1-293) x AY279380 (1-1370)
QY 1 MetAlaThrAlaArgProProTropMetTropValLeuCyAlaLeuLeuThrAlaLeu 20
DB 128 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGCTGTGTACAGCTTGGCTT 187
QY 21 LeuG1ValThrG1uH1eValLeuAlaAsnAspValSerCyAspHisProSerAn 40
DB 188 CTGGGGGTACAGAGATGTCTCCGCCAATATGATTTCTGTATACACCCCTCTAAC 247
QY 41 ThrValProSerG1ySerAsnG1nAspLeuG1yAlaG1yAlaG1yAlaAspA 60
DB 248 ACCGTCCCTCTGGAGCAACGACCTGGAGCTGGGGCCGGGAGAGCCCGGTG 307
QY 61 AspAspSerSerSerArg1le1leAsnG1ySerAspCyAspMetHisThrG1nPro 80
DB 308 GATGACAGACGACGCGCATCATCATGATCCGATGTGACACCCGCGGTG 367
QY 81 G1nAlaAlaLeuLeuLeuArgProAsnG1nLeuTyrCyG1yAlaValLeuValHisPro 100
DB 368 CAGGCCGCGCTGTGTAGAGCCCAACAGCTTACTGGGGGGGTGTGTGTGATCCA 427
QY 101 G1nTropLeuLeuThrAlaAlaHisCyAspG1yValPheArgValArgLeuG1yHis 120
DB 428 CAGTGGCTGTACAGCGCCGCCACTGCAGAGAAAGTTTTCAGATCCGCTCGGCCAC 487
QY 121 TyrSerLeuSerProValTyrG1uSerG1yG1nMetPheG1nG1yVallylsSer1le 140
DB 488 TACTCCCTGTACCAATTTATGATCTGGGACAGCATGTTCAGGGGGGTCAATCCATC 547
QY 141 ProHisProG1yTyrSerHisProG1yHisSerAspAspLeuMetLeuTyrLeuAsn 160
DB 548 CCCCACCTGGCTACTCCACCTGGCACTCTAACGACTCATGTCTCAATCACTGAAC 607
QY 161 ArgArg1leArgProThrTyrAspValArgPro1leAsnValSerSerHisCyProSer 180
DB 608 AGAAGATTCGTCCCACTAAAGATGTCAAGCCCATCAAGCTCTCTCATTTGCTCTCT 667
QY 181 AlaG1yThrLeuCyAlaLeuValSerG1yTropG1yThrThrlySerProG1nValHisPhe 200
DB 668 GCTGGGACAAAGTGTGTGTCTGGCTGGGGGACCAAGAGCCCAAGGCACTTC 727
QY 201 ProLyValLeuG1nCyLeuAsn1leSerValLeuSerG1nLyAspArgCyG1uAspAla 220
DB 728 CTTAAGCTCTCCAGTCTTGAATATCAAGCTGTCTAAGTCAAGAAAGGTGCGAGATGCT 787
QY 221 TyrProArgG1n1leAspAspThrMetPheCyAlaG1yAspLyAlaG1yArgAspSer 240
DB 788 TACCCGACAGATATATGACACCATGTTCTGGCGGGTGACAAAGCAGGTAGAGCTCC 847
QY 241 CyG1nG1yAspSerG1yG1yProValValCyAsnG1ySerLeuG1nG1yValSer 260
DB 848 TCCAGAGGTATCTGGGGGGCTGTGTGTGTGCAATGAGCTCTCTGAGAGGATCCGTGCTC 907
QY 261 TTPG1yAspTyrProCyAla1leArgProAsnArgProG1yValTyrThrAsnLeuCyAlas 280
DB 908 TGGGAGATTTACCTTTGTGTGCGGCCCAACAGACCGGGGTGTACAGAACTCTGCAAG 967
QY 281 PheThrLyStrP1leG1nG1uThr1leG1nAlaAsnSer 293
DB 968 TTCACCAAGTGGATCCAGAAACCATCCAGGCCAATCTCC 1006
RESULT 4
BD107879 1381 bp DNA linear PAT 18-SHP-2002
LOCUS BD107879 36 human secreted proteins.
DEFINITION BD107879
ACCESSION BD107879
VERSION BD107879.1 GI:23202697
KEYWORDS JP 2002500035-A/50.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1381)
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002500035-A 50 08-JAN-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002500035-A/50
PD 08-JAN-2002
PR 06-JAN-1999 JP 2000527554
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN
NI,
PI LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K39/395,A61K48/00,
PC A61P5/00,
PC A61P7/00,A61P11/06,A61P19/02,A61P29/00,A61P31/18,A61P35/00, PC
A61P35/02,
PC A61P37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC
C12N15/00,
PC A61K37/02,C12N5/00
PC A61K37/02,C12N5/00
CC 36 human secreted proteins
FH Key location/Qualifiers
FT source 1..1381
FT location/Qualifiers
FT /organism='Homo sapiens (human)'.
FEATURES
source 1..1381
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 1,296-120 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-936-271b-14 (1-293) x BD107879 (1-1381)
QY 1 MetAlaThrAlaArgProProTropMetTropValLeuCyAlaLeuLeuThrAlaLeu 20
DB 129 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGCTGTGTATCAAGCTTGGCTT 188
QY 21 LeuG1ValThrG1uH1eValLeuAlaAsnAspValSerCyAspHisProSerAn 40
DB 189 CTGGGGGTACAGACATGTTCTCCGCCAATATGATTTCTGTGACCAACCCCTCTAAC 248
QY 41 ThrValProSerG1ySerAsnG1nAspLeuG1yAlaG1yAlaG1yAlaAspA 60
DB 249 ACCGTCCCTCTGGAGCAACGACCTGGAGCTGGGGCCGGGAGAGACCCCGGTG 308
QY 61 AspAspSerSerSerArg1le1leAsnG1ySerAspCyAspMetHisThrG1nPro 80
DB 309 GATGACAGACGACGCGCATCATCATGATCCGATGTGACACCCGCGGTG 368
QY 81 G1nAlaAlaLeuLeuLeuArgProAsnG1nLeuTyrCyG1yAlaValLeuValHisPro 100
DB 369 CAGGCCGCGCTGTGTAGAGCCCAACAGCTCTACTCGGGGCGGTGTGTGATCCA 428
QY 101 G1nTropLeuLeuThrAlaAlaHisCyAspG1yValPheArgValArgLeuG1yHis 120
DB 429 CAGTGGCTGTACAGCGCCGCCACTGCAGAGAAAGTTTTCAGATCCGCTCGGCCAC 488
QY 121 TyrSerLeuSerProValTyrG1uSerG1yG1nMetPheG1nG1yVallylsSer1le 140
DB 489 TACTCCCTGTACCAATTTATGATCTGGGACAGCATGTTCAGGGGGGTCAATCCATC 548

| | | | |
|----|-----|--|------|
| QY | 141 | PRDHSAPROGLTYTYSerIaSPROGLYHSSerAaAPLeuMeLeuIIeLYaSPLeuN | 160 |
| DB | 549 | CCCCACCTGGGTACTCCACCCCTGGCACTTAACGACTATGCTATCAACTGAAC | 608 |
| QY | 161 | ARGARGLIeARgPROThrLYaSPVaIARgPROIleASnVaISerSerHiAcYbROSer | 180 |
| DB | 609 | AGAAATAATTCGTCCCACTMAAGATGTCAGACCCATCAACGTCTCTCATATGTCCCT | 668 |
| QY | 181 | AlAGIYTHrLYCyLeuVaISerGIYTPGlyTHrTHaLYSerPROGLVaIHISpHe | 200 |
| DB | 669 | GCTGGAGCAAAAGTGGCTGTGTCTGTGGGGGACAACCAAGAGCCCCCAAGTGAATTC | 728 |
| QY | 201 | PROLYaVAlleuGInCYaSPLeuASnIIeSerVaIleuSerGInLYaRGCyAGIuAaPaIa | 220 |
| DB | 729 | CCTAAGTCTCTCAAGTGGCTGAATATCAAGCGGTCTAATGACGAAAGGTCGAGGATGCT | 788 |
| QY | 221 | TYRPROARGLInIIeASpAPThrMeTPhECyAlAGIyASpLYaIAGIYARgASpSer | 240 |
| DB | 789 | TACCGGAACAAGATGATGATGACACCAATGTTCTGCGCGGTGAACAAAGCAAGTAGACATCC | 848 |
| QY | 241 | CyAGInGIyAaSPSeXGIyGIYProVaIaCYaASnGIYSerLeuGInGIYLeuVaISer | 260 |
| DB | 849 | TGCCAGGGGTGATTCGGGGGGGCTGTGGTCTGCATAGCTCCCTTCAGAGGACTCGTGTCC | 908 |
| QY | 261 | TRPGIyAaPYrProCYaAlaARgPROAaNaRGProGIYVaIYTrTHASnLeuCYaLYs | 280 |
| DB | 909 | TGGGAGATTACCTTGTGTGCCCGGCCCAACAACCGGGGTGTCTACAGCAACTCTGCAAG | 968 |
| QY | 281 | PhEThrLYrTPPIleGInIuThrIIeGInIIaASnSer | 293 |
| DB | 969 | TTACCAAGTGGATCCAGAAACATTCAGAGCCAATCC | 1007 |

| LOCUS | 1387 bp | DNA | linear | PAT 27-SEP-2004 |
|---|----------------------------|--|--------|-----------------|
| RESULT 5 | | | | |
| CQ874884 | | | | |
| DEFINITION | CQ874884 | Sequence 2 from Patent WO2004077060. | | |
| ACCESSION | CQ874884 | | | |
| VERSION | CQ874884.1 | GI:52748034 | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | Homo sapiens (human) | | | |
| REFERENCE | | | | |
| AUTHORS | 1 | Diamandis, E.P. and Petraki, C. | | |
| TITLE | | Assay for detection of renal cell carcinoma | | |
| JOURNAL | | Patent: WO 2004077060-A 2 10-SEP-2004; | | |
| FEATURES | | | | |
| source | 1..1387 | Location/Qualifiers | | |
| | /organism="Homo sapiens" | | | |
| | /mol_type="unassigned DNA" | | | |
| | /db_xref="taxon:9606" | | | |
| ORIGIN | | | | |
| Alignment Scores: | | | | |
| Pred. No.: | 1.29e-120 | Length: | 1387 | |
| Score: | 1608.00 | Matches: | 293 | |
| Percent Similarity: | 100.00% | Conservative: | 0 | |
| Best Local Similarity: | 100.00% | Mismatches: | 0 | |
| Query Match: | 100.00% | Indels: | 0 | |
| DB: | 6 | Gaps: | 0 | |
| US-09-936-271B-14 (1-293) x CQ874884 (1-1387) | | | | |
| QY | 1 | MetAlaTtRrAlaArYrPrOpRtYrMeTtRrVAlLeuCYsAlAlaLeuIlleThRAlaLeuLeu | 20 | |
| Db | 170 | ATGGCTACAGCAAGACCCCCCTGGATGTGGGCTCTGTGCTCTGATACAGCCCTTAC | 229 | |
| QY | 21 | LeuGlVAlLThGhLiVaIlleuAlaLeuAspAlaSerCYsAAspHisProSerAsn | 40 | |
| Db | 230 | CTGGGGGTACAGACACATGTTCTGCCCAACAATGATGTTTCTGTGACCACCCCTTAC | 289 | |

| | | | |
|----|------|---|------|
| QY | 41 | ThrlValProSerGIYsERsAenGIAspLeuGIyValagIyValagIyValasPAlARsER | 60 |
| Db | 290 | ACCGTGCCCTTGGAGAGAACCAAGACCTTGGAGACTTGGAGCCGGAGAGACCCCGTGCC | 349 |
| QY | 61 | AaPAsPseRseRseRATgIlelleaNGIYseRAsPcYAsPMeChIstHGIInProTrP | 80 |
| Db | 350 | GATGACAGCAGCACCGCATCATCAATGAAATCCGACTGCGATATGACACCCACCGGTGG | 409 |
| QY | 81 | GIAlaAlaIleuLeuLeuHArgProhArgInleuTYrGsgIyAlaValIleuValHisPro | 100 |
| Db | 410 | CAGGCGCGGCTGTGTCTAAGGCCCAACAGACTCTACGCGGGGGGTGTGGTCATCCA | 469 |
| QY | 101 | GIInTrPleuLeuThrAlaAlaHisCYAsArguYbValPheArgValArgLeuGIYHis | 120 |
| Db | 470 | CAGTGGCTGCTCACAGCGCGCCCATCTGAGAGAGAAAGTTTCAAGATCCGCTTGGCCAC | 529 |
| QY | 121 | TYseRLeuSeRProValTYrGIuSeRGIYGIInMeTPhEGInGIyValIYsERle | 140 |
| Db | 530 | TACTCCGTCACACAGTTTATGATCTGGGAGCAGATGTTCCAGGGGGTCCAAATCCATCC | 589 |
| QY | 141 | ProHIsProGIYTYrSeRHisProGIYHisERsAenAsPleuMeTLeuIleYsLeuAsn | 160 |
| Db | 590 | CCCCACCCCTGGCTACTCCACCCCTGGGCACCTTAACACCTCAAGCTCATCAAACTGAAC | 649 |
| QY | 161 | ArgArgIleArgProThrlYsAsPValArgProIleAsnValSerseRHisCYAsProSer | 180 |
| Db | 650 | AGAAGAAATGCTCCCATTAAGAAGTCAAGACCCATCAACGCTCTCTCATATGTCCTCT | 709 |
| QY | 181 | AlaGIYThrlYsCYsLeuValSerGIYTrpGIYThrlYsSeRProGIInValHisPhe | 200 |
| Db | 710 | GCTGGAGCAAAAGTCTTGTTGTGTCTGGCTGGGGGCAACCAAGACCCCCCAAGTCACTTC | 769 |
| QY | 201 | ProLYsValleuGIInCYsLeuAsnIleSeRValIleuSeRGIInLYbArgCYsGIuAsPAla | 220 |
| Db | 770 | CCTAAGGCTCCCACTGCTTGAATATACGCTGCTTAAGTCAAGAAAGGTGGAGAGATGCT | 829 |
| QY | 221 | TYrProARgInIleAsPAsPThMePheCYsAlaGIYAsPbAlaGIYArgAsPseR | 240 |
| Db | 830 | TACCCGAGACAGATGATGACACCATGTTCTGCCGCCGTGACAAAGAGGTAGAGACTCC | 889 |
| QY | 241 | CysGIInGIYAsPseRGIYGIYProValValCYsAsnGIYseRLeuGIInGIYLeuValSeR | 260 |
| Db | 890 | TGCCAGGGGTGATTTCTGGGGGGGCGTGTGTCTGCAATGGCTCCCTGCAAGGAGCTGTGTCC | 949 |
| QY | 261 | TrpGIYAsPTrYProCYsAlaArgProAsnArgProGIYValTYrThrAsnLeuCYsArg | 280 |
| Db | 950 | TGGGGAAGATTACCTTGTGTGCCGCCCAACAGACCGGGGTGTCTACAGAAACTCTGCAAG | 1009 |
| QY | 281 | PherThrlYsTrpIleGIInGIuThrlleGIAlaIAsnSeR | 293 |
| Db | 1010 | TTTCACCAAGTGGATTCAGAAACCATCAGGCCCACTTCC | 1048 |

| RESULT | 6 | | | | |
|------------|---|-------------|-----|--------|-----------------|
| LOCUS | CO874960 | 1387 bp | DNA | linear | PAT 27-SEP-2004 |
| DEFINITION | Sequence 4 from Patent WO2004075713. | | | | |
| ACCESSION | CO874960 | | | | |
| VERSION | CO874960.1 | GI:52748059 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Homo sapiens (human) | | | | |
| | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| REFERENCE | 1 | | | | |
| AUTHORS | Diamondis, E.P. | | | | |
| TITLE | Multiple marker assay for detection of ovarian cancer | | | | |
| JOURNAL | Patent: WO 2004075713-A 4 10-SEP-2004; | | | | |
| | MOUNT SINAI HOSPITAL CORPORATION (CA) | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..1387 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="unassigned DNA" | | | | |
| | /db_xref="taxon:9606" | | | | |

ORIGIN

Alignment Scores:

| Pred. No.: | 1 296-120 | Length: | 1387 |
|------------------------|-----------|---------------|------|
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x C0874960 (1-1387)

```

QY      1 MetAlaThrAlaArgProTfPmEtrPValLeuCYsAlaLeuIleThraAlaLeuLeu 20
DB      170 ATGGCTACGACGAGACCCCTCGATGTGGGTCTGTGCTCTGATCAAGCTTGGCTT 229

QY      21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB      230 CTGGGGGTCAAGAGATGTTCTCGCAACATGATGTTCTGTGACCACTCTTAAAC 289

QY      41 ThrValProSerGlySerAsnGlnAspLeuGlyValAGlyValAGlyValAspAlaArgSer 60
DB      290 ACCGTCCTCTGGGAGCAACGAGCTGGGAGCTGGGAGCGGAGAGAGCGCCGGTGG 349

QY      61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTfP 80
DB      350 GATGACAGCAGAGCGCGCATCATCATGATGATCGGACTGGCATATGCACACCGCGGTG 409

QY      81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB      410 CAGGCGGCGCTGTGGTAAAGCCCAACAGCTTACTGGGGCGGTGTGGTGCATCA 469

QY      101 GlnTfPLeuLeuThraAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
DB      470 CAGTGGCTGTCAAGCGCCGCCCACTCAAGAGAAAGTTTCAAGATCCGTCTCGGCCAC 529

QY      121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValYsSerIle 140
DB      530 TACTCCCTGTCAACAGTTATGAAATCTGGGACAGATGTTCCAGGGGGTCAAAATCCATC 589

QY      141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleYsLeuAsn 160
DB      590 CCCACCTGGCTACCTCCACCTGGGCACTTAAAGACCTCAATGCTCAATCAAACTGAAC 649

QY      161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB      650 AGAAGATTCGTCCCACTAAAGATGCAAGCCCATCAAGCTCTCTCTCATTTGCTCTCT 709

QY      181 AlaGlyThrIleYsLeuValSerGlyTfPGLyThrThrIleSerProGlnValHisPhe 200
DB      710 GCTGGGACAAAGTCTTGGGTGTGGCTGGGGGCAACCAAGAGGCCCCCAAGTGCATTC 769

QY      201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleYsArgCysGluAspAla 220
DB      770 CCTAAAGTCTCTCAAGCTTGAATATCAAGCTGCTTAAAGTCAAGAAAGTGCAGAGTGT 829

QY      221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValIleArgAspSer 240
DB      830 TACCCAGACAGATGATGACACCATGTTCTGCGCGGAGACAAAGAGGATAAGACTCC 889

QY      241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB      890 TGCACAGGTGATTTCTGGGGGGCTGTGTGTGCAATGGGTCCCTGCAGGGAGCTCGTTC 949

QY      261 TfPGLyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCYsIle 280
DB      950 TGGGAGATTAACCTTGTGCTGCGGCCCAACAGACCGGAGTCTTACACGAACCTCTCAAG 1009

QY      281 PheThrIleTfPLeuGlnGlnIleGlnIleGlnAlaAsnSer 293
DB      1010 TTCACCAAGTGAATCCAGGAACCATCCAGGCCAATCC 1048
  
```

RESULT 7

AF168768 1387 bp mRNA linear PRI 16-Apr-2002
 LOCUS Homo sapiens stratum corneum tryptic enzyme (SCTE) mRNA, complete
 DEFINITION cds
 ACCESSION AF168768
 VERSION AF168768.2 GI:20153423
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation
 J. Biol. Chem. 274 (42), 30033-30040 (1999)

JOURNAL
 MEDLINE
 PUBMED 99445563
 10514489

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE Direct Submission
 Submitted (12-JUL-1999) Public Health and Clinical Medicine,
 Dermatology and Venereology, Umea University, University Hospital,
 Umea SE-901 85, Sweden
 3 (bases 1 to 1387)
 Brattand, M. and Egelrud, T.
 Direct Submission

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE Submitted (16-Apr-2002) Public Health and Clinical Medicine,
 Dermatology and Venereology, Umea University, University Hospital,
 Umea SE-901 85, Sweden
 Sequence update by submitter
 On Apr 16, 2002 this sequence version replaced gi:6063032.

REMARK
 COMMENT location/Qualifiers
 FEATURES
 source

1..1387
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="stratum corneum"
 1..1387
 /gene="SCTE"
 170..1051
 /gene="SCTE"
 /note="SCTE, trypsin-like serine protease; possibly involved in desquamation"
 /codon_start=1
 /product="stratum corneum tryptic enzyme"
 /protein_id="AA03101.1"
 /db_xref="GI:6063033"
 /translation="MATARPMPWVLCALITALLGVTEHYLVANNVDSCHPSNTVPS
 GSNODLAGAGEDARSDSSRIINGSDCHTQMOALILRPOLYCGAVLHVQW
 LITTAHCKRVPRVRLIGHYSLSPVSEGOQMOGKSLIPPEYSHRGSNDMLTKLN
 RIRPTKVRPRINVSHPKSLKLVSGMKTKSPQVHFPRVLOCLINISVSQKCE
 DAVPRQIDITMECADDKGRDSQGDSPVVCNSLDELVSMDYPCARPNRPVYT
 NICKETKWIQETIQANS"

ORIGIN

Alignment Scores:

| Pred. No.: | 1 296-120 | Length: | 1387 |
|------------------------|-----------|---------------|------|
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x AF168768 (1-1387)

```

QY      1 MetAlaThrAlaArgProTfPmEtrPValLeuCYsAlaLeuIleThraAlaLeuLeu 20
DB      170 ATGGCTACGACGAGACCCCTCGATGTGGGTCTGTGCTCTGATCAAGCTTGGCTT 229

QY      21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
  
```


Db 230 CTGGGGGTCAAGAGCATGTTCTGCCCAACATGATGTTCTGTGACACCCCTCTAAC 289
Qy 41 ThTValProSerGlySerAsnGlnAspLeuGlyValAGlyValAGlyValAspSer 60
Db 290 ACCGTGCTCTTGGAGCAACACAGACCTGGGAGCTGGGGCGGGAAGAGCCGGTGG 349
Qy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 350 GATGACAGCAGACAGCCGCACTCATATGATCCAGCTGCATATGCACACCCAGCCGTGG 409
Qy 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValAlaLeuValHisPro 100
Db 410 CAGGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGGGGGGGGGTGGTGATCCA 469
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgValysValPheArgValArgLeuGlyHis 120
Db 470 CAGTGGCTGTCTACAGCGCGCCGCTGACAGGAAGATTTTCAAGTCCGCTCGGGCAC 529
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerTle 140
Db 530 TACTCCCTGTCCACAGTTTATGATCTGGGCGAGATGTTCCAGGGGGTCAATCCATC 589
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleTyrLeuAsn 160
Db 590 CCCACCTGGCTACTCCACCTGGCCACTGACGAAAGATTTTCAAGTCCGCTCGGGCAC 649
Qy 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
Db 650 AGAAGATTCGTCCCTCAAGATGATGAGACCCATCAAGCTCTCTCTCATTTGCTCTCT 709
Qy 181 AlaGlyThrIleCysLeuValSerGlyTyrGlyTyrThrIleCysSerProGlnValHisPhe 200
Db 710 GCTGGGCAAAAGTCTTGGTGTCTGGCTGGGGGCAACCAAGACCCCAAGTCACTTC 769
Qy 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
Db 770 CTTAAGGCTCTCCAGTCTTGAATATCAGCGCTGTAAGTGAAGAGGTGGAGAGTCT 829
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyValAspValAGlyValAspSer 240
Db 830 TACCCGAGACAGATGATGACACCATGTTCTGCCCGGTGCAAAAGAGGTAGAGACTCC 889
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 890 TGCAGGGGTATCTGGGGGGCTGTGTCTGCAATGCTCTCCAGGAGACTGTCTCTC 949
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 950 TGGGAGATTAACCTTGTGCCCCGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1009
Qy 281 PheThrIleTyrTrpIleGlnGlyThrIleGlnAlaAsnSer 293
Db 1010 TTCACCAAGTGCATCCAGAAACCATCCAGGCCCACTCC 1048

RESULT 8
AY279381 1438 bp mRNA linear PRI 26-MAY-2003
LOCUS Homo sapiens kallikrein 5 splice variant 2 (KLKS) mRNA, complete
DEFINITION cds; alternatively spliced.
ACCESSION AY279381
VERSION AY279381.1 GI:31075482
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1438)
AUTHORS Kurlender, L., Yousef, G.M., White, N.M.A., Robb, J.-D., Borgono, C.A.
and Diamandis, E.P.
TITLE Identification of splice variants for the human kallikrein gene 5
(KLKS)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1438)

AUTHORS Yousef, G.M., Robb, J.-D., White, N.M.A., Kurlender, L., Borgono, C.A.
and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
source
1. 1438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. 1438
/gene="KLKS"
1. 195
/gene="KLKS"
196. 1077
/gene="KLKS"
/note="serine protease; differentially expressed in
malignancy; may play a role in desquamation of skin; hks;
alternatively spliced; similar to the product of Genbank
Accession Number AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 2"
/protein_id="AAP42276.1"
/db_xref="GI:31075483"
/translation="MATARPMMWVLCALITALLGVTEHTVLANVDSCPSNTVPS
GSNQDAGAGGEBARDSDSSRIINGSDCMHPQALLLRPNOLYCAVIVHPQM
LTPAARCKKVPFVRLGHSLSPEVSSGQMPQGVKSIPIHGYSHPSHNDMLIKLN
RRIRPTKDVPIVSSHCPAGTKCLVSGMTKSPVHPKYLQCLNIVLSOKRCE
DAYRQIDPTMPCAGKAGRDSQGSBSGSEVYVNGSLQGLVSGDYPCARPNNPGVYT
NICKFTKWIOETIQANS"
3' UTR
1078. 1438
/gene="KLKS"
1387. 1392
polya_signal
/gene="KLKS"
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-120 Length: 1438
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-936-271b-14 (1-293) x AY279381 (1-1438)
Qy 1 MetAlaThrAlaArgProPOTtrpMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
Db 196 ATGGCTACAGCAAGACCCCTGATGTGGTCTCTGTGATCAGAGCTTGGCTT 255
Qy 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 256 CTGGGGGTCAAGAGCATGTTCTGCCCAACATGATGTTCTCTGTACCAACCCCTCTAC 315
Qy 41 ThTValProSerGlySerAsnGlnAspLeuGlyValAGlyValAGlyValAspSer 60
Db 316 ACCGTGCTCTTGGAGCAACACAGACCTGGAGCTGGGGCGGGAAGAGCCGGTGG 375
Qy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 376 GATGACAGCAGACAGCCGCACTCATATGATCCAGCTGCATATGCACACCCAGCCGTGG 435
Qy 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValAlaLeuValHisPro 100
Db 436 CAGGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGGGGGGGGGTGGTGATCCA 495
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgValysValPheArgValArgLeuGlyHis 120
Db 496 CAGTGGCTGTCTACAGCGCGCCGCTGACAGGAAGATTTTCAAGTCCGCTCGGGCAC 555

QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIlySerIle 140
Db 556 TACTCCCTGTCAACAGTATTATGATCTGGGACAGACATGTTCCAGGGGGTCAAAATCCATC 615
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleIlyValLeuAsn 160
Db 616 CCCACCCCGGTGACTCCACCCCTGGGCACTCTAACGACCTCATGTCTCAACAACTGAAC 675
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 676 AGAAGAAATTCGTCCCACTTAAGATGTCAAGCCATCAAGCTCTCTCATTTGTCCCTCT 735
QY 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrThrIlySerProGlnValHisPhe 200
Db 736 GCTGGGACAAGGCTTGCTGTGCTGGGAGCAACCAAGAGCCCCCAAGTGCACTTC 795
QY 201 ProIlyValIleuGlnIlyLeuAsnIleSerValLeuSerGlnIlyArgCysGlyIlyAspAla 220
Db 796 CCTAAGGTCTCCAGCTTGAATATCAGCTGTCAAGTCAAGAAAGGTGCGAGATGCT 855
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyValArgAspSer 240
Db 856 TACCCAGACAGATATATGACACCATGTTCTGCGCCGTGACAAAGCAGGTAGACATCC 915
QY 241 CysGlnGlyAspSerGlyGlyProValIlyCysAsnGlySerLeuGlnIlyLeuValSer 260
Db 916 TGCACAGGTGATCTGGGGGGCTGTGTCTGTGCAATGGCTCCCTGCAGGGAGCTCGTCTC 975
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlyAs 280
Db 976 TGGGAGATTAACCTTGTCCTGCGCCGACCAACAGCCGGGTGTCAACGAACCTCTGCAG 1035
QY 281 PheThrIlyStrIleGlnIlyThrIleGlnIlyAsnSer 293
Db 1036 TTCACCAAGTGATCCAGGAACCATCCAGGCCAATCTCC 1074

RESULT 9
LOCUS AR352504 1499 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 1 from patent US 6589770.
ACCESSION AR352504
VERSION AR352504.1 GI:33757710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1499)
AUTHORS Kitado,H., Yoshikawa,A. and Zaiki,T.
TITLE Keratinocyte derived protease
JOURNAL Patent: US 6589770-A 1 08-JUL-2003;
FEATURES
source 1..1499
1..1499
/organism="unknown"
/mol_type="genomic DNA"
Alignment Scores:
Pred. No.: 1..4e-120 Length: 1499
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-936-271B-14 (1-293) x AR352504 (1-1499)

QY 1 MetAlaThrAlaArgProProTyrMetTyrValIleuCysAlaIleuIleThrAlaIleuLeu 20
Db 291 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTCTGTGCTCTGATCAAGCTTGCTT 350
QY 21 LeuGlyValThrGlnHisValIleuAlaAsnArgValSerCysAspHisProSerAsn 40
Db 351 CTGGGGGTCAAGAGCATGTTCTCGCCAACAATGATGTTCTGTGACCAACCCCTCTAAC 410

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValIlyGlyIlyAspAlaArgSer 60
Db 411 ACCGTGCTCTGTGGAGGACACAGAGACTGGAGACTGGGGCCGGGGAAGACCCCGGTG 470
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80
Db 471 GATGACAGCAGCAGCCGATCATCATGGATCCGACTGGATATGACACCCAGCGCTGG 530
QY 81 GlnAlaAlaLeuLeuMetArgProAsnGlnLeuTyrCysGlyValValLeuValHisPro 100
Db 531 CAGGCGGCTGTCTTGAAGCCCAACAGCTTACTGTGGGGCGGTGTGGATGCCA 590
QY 101 GlnTyrPleuLeuThrAlaAlaHisCysArgIlyIlyValPheArgValArgLeuGlyHis 120
Db 591 CAGTGGCTGTCAACGGCCGACCTGACAGGAAGAAAGTTTCAAGTCCGTCTCGGCCAC 650
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIlySerIle 140
Db 651 TACTCCCTGTCAACAGTTTATGATATCTGGGACAGATGTTCCAGGGGGTCAAAATCCATC 710
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleIlyLeuAsn 160
Db 711 CCCACCTGCTGACTCCCACTGACCTTAACGACCTTATGCTCATCAACTGAAC 770
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 771 AGAAGAAATTCGTCCCACTTAAGATGTCAAGCCCATCAAGCTCTCTCATTTGTCCCTCT 830
QY 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrThrIlySerProGlnValHisPhe 200
Db 831 GCTGGGACAAGGTGCTTGCTGTCTGCTGGGGGACCAACAGAGCCCAAGTGCATCTTC 890
QY 201 ProIlyValIleuGlnCysLeuAsnIleSerValLeuSerGlnIlyArgCysGlyIlyAspAla 220
Db 891 CCTAAGGTCTCCAGGTCTGAATATCAGCGTGTCAAGTCAAGAAAGGTGAGAGATGCT 950
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyValArgAspSer 240
Db 951 TACCCAGACAGATATGATGACACCATGTTCTGCGCCGTGACAAAGCAGGTAGACATCC 1010
QY 241 CysGlnGlyAspSerGlyGlyProValIlyCysAsnGlySerLeuGlnIlyLeuValSer 260
Db 1011 TGCACAGGTGATCTGGGGGGCTGTGTGTGCAATGCTCCCTGCAGGGAGCTCGTCTCC 1070
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlyAs 280
Db 1071 TGGGAGATTAACCTTGTCCTGCGCCGACCAACAGCCGGGTGTCAACGAACTCTGCAG 1130
QY 281 PheThrIlyStrIleGlnIlyThrIleGlnIlyAsnSer 293
Db 1131 TTCACCAAGTGATCCAGGAACCATCCAGGCCAATCTCC 1169

RESULT 10
LOCUS BD005362 1499 bp DNA linear PAT 31-JAN-2002
DEFINITION Protease.
ACCESSION BD005362
VERSION BD005362.1 GI:18633733
KEYWORDS JP 2001501837-A/1.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1499)
AUTHORS Kitado,H., Yoshikawa,A. and Zaiki,T.
TITLE Protease
JOURNAL Patent: JP 2001501837-A 1 13-FEB-2001;
COMMENT THE PROCTER & GAMBLE CO
OS Human keratinocyte
PN JP 2001501837-A/1
PD 13-FEB-2001
PR 03-OCT-1997 JP 1999521662

PI HARUO KITADO, AKIKAZU YOSHIKAWA, TOMOKO ZAIKI
PC C12N15/57, C12N9/64, A61K38/48, A61K7/48, C11D3/386, C07K16/40 CC

PH Key Location/Qualifiers
FT CDS (291) . (1172)
FT mat peptide (489) . (1172)
FT 5' UTR (1) . (290)
FT 3' UTR (1173) . (1499).

FEATURES
source
1. .1499
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 1.4e-120 Length: 1499
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x BD005362 (1-1499)

OY 1 MetAlaThrAlaArgProTpmETrpValLeuCyAlaLeuIleThraAlaLeu 20
DB 291 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTCTCTGATCACAGCCCTGCTT 350
OY 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 351 CTGGGGGTCAAGAGCATGTTCTGCCAACATGATGTTCTGTGACCAACCCCTCTAAC 410
OY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValLeuAlaArgSer 60
DB 411 ACCGTCCTCTGGGAGCAACAGACCTGGAGCTGGGCGGGAGAGACGCCGCTGC 470
OY 61 AsnAspSerSerArgGlyIleLeuAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 471 GATGACAGCAGCAGCCGATCATCAAGATCCGACTGCAATATGACACCCAGCCGTGG 530
OY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCyAspGlyAlaValLeuValHisPro 100
DB 531 CAGGCGCGCTGTGCTACAGGCCCAACCACTACTGCGGGCGGTGTGTCATCCA 590
OY 101 GlnTrpLeuLeuThraAlaAlaHisCyAspGlyValValPheArgValAlaGlyLeuGlyHis 120
DB 591 CAGTGGCTGCTCAAGCGCGCCCACTGCAGAGAGAAAGTTTTCAGAGTCCGCTCGGCAC 650
OY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLeuValSerIle 140
DB 651 TACTCCCTGTCCACAGATTATGATCTGGGAGAGATGTTCCAGGGGGTCAAAATTCATC 710
OY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleTyrLeuAsn 160
DB 711 CCCCAACCTGGTACTCCACCTGGCCACTCTAACACCTCATGCTCAATCAACTAAC 770
OY 161 ArgArgIleArgProThrIlyAspValAlaArgProIleAsnValSerSerHisCyAspProSer 180
DB 771 AGAAGAAATTCGTCCCACTAAAGATGTCAGACCCATCAACGTCCTCTCATTTGCCCTCT 830
OY 181 AlaGlyThrIlySerCyAlaLeuValSerGlyTyrGlyThrThrIlySerProGlnValHisPhe 200
DB 831 GCTGGGCAAGAGCTGTGTCTGGCTGGGGCAACCAAGAGCCCAAGAGTGCATCTTC 890
OY 201 ProIlyValLeuGlnCyAlaAsnIleSerValLeuSerGlnIlyAspCyAspGlyAspAla 220
DB 891 CTTAAAGTCTCTCAAGTCTTGAATATCAAGCTGTGTAAGTCAAGAAAGTGGAGAGTCTT 950
OY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyValAspIlyAlaGlyValArgAspSer 240
DB 951 TACCCGAGACAGATGATGACACCATGTTCTGCGCCGGTGAACAAAGCAGTAGAGATCC 1010

OY 241 CyGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 1011 TGGCAAGGATTCCTGGGGGGCTGTGTGCTGAATGGCTCCCTGCAGGAGCTCGTCTC 1070
OY 261 TrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyGly 280
DB 1071 TGGGAGATTACCTTCCTTGTGCGCCGCAACAGACCCGGTGTCAACGAACCTCTCAAG 1130
OY 281 PheThrIleTrpIleGlnGlnIleGlnIleGlnAlaAsnSer 293
DB 1131 TTCACCAAGTGAATCCAGAAACCATTCAGGCCACTCC 1169

RESULT 11
BD107865
LOCUS
DEFINITION 36 human secreted proteins. 1516 bp DNA linear PAT 18-SEP-2002
BD107865
ACCESSION
VERSION BD107865.1 GI:23202683
KEYWORDS JP 2002500035-A/36.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1516)
Ruben, S.M., Soppet, D.R., Ebner, R., Lafleur, D.W., Ni, J.,
Brewer, L.A., Olsen, H.S., Duan, R.D. and Rosen, C.A.
36 human secreted proteins
Patent: JP 2002500035-A 36 08-JAN-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002500035-A/36
PD 08-JAN-2002
PR 06-JAN-1999 JP 2000527554
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN
NI,
PI LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K39/395, A61K48/00,
PC A61P5/00, A61P7/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC
A61P35/02,
PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC
C12N15/00,
PC A61K37/02, C12N5/00
CC n equals a, c, g, or c
FH Key Location/Qualifiers
FT source 1. .1516
/organism="Homo sapiens (human)".
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1. .1516
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.42e-120 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x BD107865 (1-1516)

OY 1 MetAlaThrAlaArgProTpmETrpValLeuCyAlaLeuIleThraAlaLeu 20
DB 254 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTCTCTGATCACAGCCCTGCTT 313
OY 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 314 CTGGGGGTCAAGAGCATGTTCTGCCAACATGATGTTTCTGTGACCAACCCCTCTAAC 373

QY 41 ThrValProSerGlySerAnglnAapLeuGlyAlaGlyValAlaGlyIleuAapAlaArgSer 60
 Db 374 ACCGTCCTCTCTGGAGCAACCAAGACCTGGAGAGCTGGGGCCGGGAAGACGCCCGGTG 433
 QY 61 AspAapSerSerSerArgIleIleAngIySerAapCyAapMethIsthrgInProTrp 80
 Db 434 GATGACAGACAGACCGGCATCATCATGATGATCGACTGCGATATGACACCCAGCGGTG 493
 QY 81 GlnAlaAlaLeuLeuLeuAapProAanglnLeuTyCyGlyValAlaValLeuValHisPro 100
 Db 494 CAGGCGCGCTGTGTCTAAGCCCAACCAAGCTTACTGCGGGCGGTGTGGTCATCCA 553
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAapGlyValPheArgValArgLeuGlyHis 120
 Db 554 CAGTGGCTGTCAACGCGCCCACTGACAGAAAGATTTCAGATCCGTCTCGGCCAC 613
 QY 121 TyrSerLeuSerProValTyrgIuSerGlyGlnIleMecPheGlnGlyValIySerIle 140
 Db 614 TACTCCCTGTCAACAGTTTATGATCTGGGAGCAGATGTTCCAGGGGTCATATCCATC 673
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAapAapLeuMethLeuIleValLeuAn 160
 Db 674 CCCACCTGGCTACTCCACCTGGCCACTTACAGACTCTATGCTCATCAAACTGAAC 733
 QY 161 ArgArgIleArgProThrIyAapValArgProIleAapValSerSerHisCyProSer 180
 Db 734 AGAAGAAATTCGTCCTCAAGATGTCAGACCATCACTCTCTCTATCTCTCTCTCT 793
 QY 181 AlaGlyThrIyGlySerValSerGlyTyrGlyTyrThrIySerProGlnValHisPhe 200
 Db 794 GCTGGGACAAAGCTTGTGTGCTGGCTGGGAGCAACCAAGAGCCCAAGTGCACCTTC 853
 QY 201 ProIyValIleGlnIyCysLeuAnIleSerValIleSerGlnIyAapGlyValAapAla 220
 Db 854 CTTAAGCTCTCCAGTGTCTGAATATCAGGTGCTAAGTCAAGAAAGTGCAGGATGCT 913
 QY 221 TyrProArgGlnIleAapAapThrMetPheCyAalaGlyAapIyValArgAapSer 240
 Db 914 TACCCGACAGATATGACACCATGTTCTGCGCGGTGACAAACAGTAGTAGCTCC 973
 QY 241 CyGlnGlyAapSerGlyIyProValIyCyAangIySerLeuGlnIyLeuValSer 260
 Db 974 TCCACAGGCTGATCTGGGGGCTGTGTGCTGCAAGGCTCCCTGACAGGACTCGATCC 1033
 QY 261 TrpGlyAapTrpProCyAalaArgProAangIyValTyrThrLeuLeuCyIyVal 280
 Db 1034 TGGGGGATTAACCTTGTGCTGCGGCCCAACAGCCGGGTGTCTACAGAACTTGTCAAG 1093
 QY 281 PheThrIyTrpIleGlnIyThrIleGlnAlaAangSer 293
 Db 1094 TTCACCAAGTGTATCCAGAAACCATCCAGGCCAATCC 1132
 RESULT 12
 LOCUS BC008036 1527 bp mRNA linear PRI 29-JUN-2004
 DEFINITION Homo sapiens kallikrein 5, mRNA (cDNA clone MGC:1107
 IMAGE:2989806), complete cde.
 ACCESSION BC008036
 VERSION BC008036.2 GI:33873788
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1527)
 AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stajich, M.F., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullan, S.J., Bosak, S.A., McSwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S.,
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smal, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1527)
 Strausberg, R.
 Direct Submission
 Submitted (21-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:14165477.
 CONTACT: MGC help desk
 Email: cgabs-rt@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadn@systemsbio.org
 Anup Madan, Jessica Fahy, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAL Plate: 3 Row: m Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 22208993.
 FEATURES
 source
 1..1527
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:1107 IMAGE:2989806"
 /issue_type="Ovary, adenocarcinoma"
 /clone_1fb="NIH MGC_9"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1527
 /gene="KLK5"
 /note="synonym: SCTB, KLK-L2, KLKL2"
 /db_xref="locusid:25818"
 /db_xref="MIM:605643"
 285..1166
 /gene="KLK5"
 /codon_start=1
 /product="kallikrein 5, preproprotein"
 /protein_id="AAH08036.1"
 /db_xref="GI:14165478"
 /db_xref="locusid:25818"
 /db_xref="MIM:605643"
 /translation="MATHRPFPMWVLCALITRLILGTEHYLANNDVSCDHSNTVPS
 GSNDDLAGAGSDADSDSSRTINGSDCDHTDQALILRPQLTGAVALVHPQW
 LITAAHCKKVFRLVILGHVSLSPVYSGQWFOGVKSIPHEVSHRGSNDLMLTKLN
 RIRPRTKVRILNVSSHCPASGATGLVSGWGTTSPOVHPFVTLQCLNISVLSQKRC
 DAYPQIDDTMECAQDKGRDSCQSDSGPVVNCNGLQGLVSWGVPQARFRRPEVYT
 NLCKRTKRIQETIQANS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.43e-120 Length: 1527

Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x BC008036 (1-1527)

QY 1 MetAlaThrAlaArgProTpmETpValLeuCyAlaLeuIleThrAlaLeu 20
 Db 285 ATGGCTTACAGAACACCCCTGATGTGGTGTCTGTCTGATCAGACCTTCTT 344
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
 Db 345 CTGGGGGTTCACAGACATGTTCTGCCCAATATGTTCTTGTGACCACTTCTTAC 404
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAspAlaArgSer 60
 Db 405 ACCGTGCTCTGGAGACACACAGACTGGGACTGGGCGGGGAAGACCGCGTGG 464
 QY 61 AspAspSerSerSerAlaGlyIleIleAsnGlySerAspCyAspMetHisThrGlnProTpm 80
 Db 465 GATGACAGACAGACCCGATCATCATGATCCGATCGCATATGACACCCACCGCTGG 524
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyAlaValLeuValHisPro 100
 Db 525 CAGCCCGCTGTTGCTTACAGCCCAACAGCTTACGCGGGGCGGTGTGTGATCCA 584
 QY 101 GlnTpmLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
 Db 585 CAGTGGCTGTCTCAGCGCGCCCACTGACAGAAAGTTTTCAGAGTCCGCTCGGCAC 644
 QY 121 TyrsSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
 Db 645 TACTCCCTGTCCACAGATTATGAATCTGGGACAGATGTTCCAGGGGCTCAAAATCCATC 704
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValLeuAsn 160
 Db 705 CCCCACCTGTGCTACTCCACCTGGCCACTCTTAAAGCTCATAGCTCATTAACCTGAC 764
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyProSer 180
 Db 765 AGAAGATTGCTCCCACTAAAGATGACAGACCATCAACGCTCTCTCATTTGCCCTCT 824
 QY 181 AlaGlyThrIleValCyLeuValSerGlyTyrGlyThrThrIleValSerProGlnValHisPhe 200
 Db 825 GCTGGGACAAAGTCTGTGTCTGTGGCTGGGGACAAACCAAGACCCCAAGTGCACCTTC 884
 QY 201 ProIleValLeuGlnCyLeuValHisValLeuSerGlnIleValArgCyGluAspAla 220
 Db 885 CTTAAGGCTCTCCAGTGTCTTAATATCAGCGTGTCTAAGTCAAAAGGTGGAGAGTGT 944
 QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAlaGlyArgAspSer 240
 Db 945 TACCGAGACAGATAGATGACACCATGTTTGGCGCGGTGACAAAGCAGTGAAGACTCC 1004
 QY 241 CyseGlnIleAspSerGlyIleProValValCyAsnGlySerLeuGlnGlyValIser 260
 Db 1005 TGCCAGGCTGATTTCTGGGGGCGCTGTGTCTGCAATGCTCTCCGACAGGACTGTGCTCC 1064
 QY 261 ThrGlyAspTyrProCyAlaArgProAspArgProGlyValTyrThrAsnLeuCyValys 280
 Db 1065 TGGGAGATTAACCTTGTGTGGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAG 1124
 QY 281 PheThrIleValIleGlnGluThrIleGlnAlaAsnSer 293
 Db 1125 TTCAACCAAGTGCATCCAGAAACCATCCAGGCCAACTCC 1163

RESULT 13
 AR525280 AR525280 1570 bp DNA linear PAT 20-DEC-2002
 LOCUS AR525280
 DEFINITION Sequence 308 from patent US 6478825.
 ACCESSION AR525280

VERSION AR525280.1 GI:27300488
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1570)
 AUTHORS Winterbottom,J.M., Shimp,L., Boyce,T.M. and Kaes,D.
 TITLE Implant, method of making same and use of the implant for the
 treatment of bone defects
 JOURNAL Patent: US 6478825-A 308 12-NOV-2002;
 FEATURES Location/Qualifiers
 source 1..1570
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,48e-120 Length: 1570
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-936-271b-14 (1-293) x AR525280 (1-1570)

QY 1 MetAlaThrAlaArgProTpmETpValLeuCyAlaLeuIleThrAlaLeu 20
 Db 354 ATGGCTTACAGAACACCCCTGATGTGGTGTCTGTCTGATCAGACCTTCTT 413
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
 Db 414 CTGGGGGTTCACAGACATGTTCTGCCCAATATGTTCTTGTGACCACTTCTTAC 473
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAspAlaArgSer 60
 Db 474 ACCGTGCTCTGGAGACACACAGACTGGGACTGGGCGGGGAAGACCGCGTGG 533
 QY 61 AspAspSerSerSerAlaGlyIleIleAsnGlySerAspCyAspMetHisThrGlnProTpm 80
 Db 534 GATGACAGACAGACCCGATCATCATGATCCGATGATGACACCCACCGCTGG 593
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyAlaValLeuValHisPro 100
 Db 594 CAGCCCGCTGTTGCTTACAGCCCAACAGCTTACGCGGGGCGGTGTGTGATCCA 653
 QY 101 GlnTpmLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
 Db 654 CAGTGGCTGTCTCAGCGCGCCCACTGACAGAAAGTTTTCAGAGTCCGCTCGGCAC 713
 QY 121 TyrsSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
 Db 714 TACTCCCTGTCCACAGATTATGAATCTGGGACACAGATGTTCCAGGGGCTCAAAATCCATC 773
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValLeuAsn 160
 Db 774 CCCCACCTGTGCTACTCCACCTGGCCACTTCAACCACTTCACTGCTCAAACTGAC 833
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyProSer 180
 Db 834 AGAAGATTGCTCCCACTAAAGATGACAGACCATCAACGCTCTCTCATTTGCCCTCT 893
 QY 181 AlaGlyThrIleValCyLeuValSerGlyTyrGlyThrThrIleValSerProGlnValHisPhe 200
 Db 894 GCTGGGACAAAGTCTGTGTGGCTGGGGACAAACCAAGGCCCCCAAGTGCACCTTC 953
 QY 201 ProIleValLeuGlnCyLeuValHisValLeuSerGlnIleValArgCyGluAspAla 220
 Db 954 CTTAAGGCTCTCCAGTGTCTTAATATCAGCGTGTCTAAGTCAAAAGGTGGAGATGTCT 1013
 QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAlaGlyArgAspSer 240
 Db 1014 TACCGAGACAGATAGATGACACCATGTTTCTGGCGCGGTGACAAAGCAGTGAAGACTCC 1073

| | | | |
|------------|---------------------------------------|---|-----------------|
| Qy | 241 | CysGInGInyAspSerSerGInyIProValyAlaCysAnGInySerLeuGInyIleuValSer | 260 |
| Db | 1074 | TGCCAGGGGTGATCTCGGGGGGCGCTGTGGTCTGCATGGCTCTCCGCAAGGACCTGGTGTCC | 1133 |
| Qy | 261 | TrpGInyAspIlyrProCysAlaArgProAnaRgpProGInyValIlyrThrAsnLeuCysIlys | 280 |
| Db | 1134 | TGGGAGATTTACCTTGTGTGCGCCGGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAG | 1193 |
| Qy | 281 | PheThrIlyrSTPileGInyIThrIleGInAlaAnSer | 293 |
| Db | 1194 | TTCAACCAAGTGGATCCAGAAACCATCCAGGCCAACTCC | 1232 |
| RESULT 14 | | | |
| LOCUS | AR528679 | 1570 bp | DNA |
| DEFINITION | Sequence 455 from patent US 6725730. | linear | PAT 08-OCT-2004 |
| ACCESSION | AR528679 | | |
| VERSION | AR528679.1 | GI:53916757 | |
| KEYWORDS | . | | |
| SOURCE | Unknown. | | |
| ORGANISM | Unknown. | | |
| REFERENCE | Unclassified. | | |
| AUTHORS | 1 (bases 1 to 1570) | | |
| TITLE | Bollinger,C.L., Jr. | | |
| JOURNAL | Crane test weight assembly and method | | |
| FEATURES | Patent: US 6725730-A 455 27-ARR-2004; | | |
| source | 1..1570 | | |

```

/organism="unknown"
/mol_type="genomic DNA"

```

ORIGIN

| | |
|------------------------|-----------|
| Alignment Scores: | |
| Pred. No.: | 1,48e-120 |
| Score: | 1608.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 100.00% |
| DB: | 6 |
| Length: | 1576 |
| Matches: | 293 |
| Conservative: | 0 |
| Mismatches: | 0 |
| Indels: | 0 |
| Gaps: | 0 |

US-09-936-271B-14 (1-293) x AR528679 (1-1570)

| | | |
|----|---|-----|
| Qy | MetAlaThrLysArgProProlTrpMetCysValLeuLeuLeuThrAlaLeuLeu | 20 |
| Dd | ATGCTACACACAGACCCCCCTGAGATGGGAGTCTGTGCTCTGATACACACCTTGCT | 411 |
| Qy | LeuGlyValThrGlnHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn | 40 |
| Dd | CTGGGGGTTCACAGAGCATGTCTTCGCCCAACAAAGATGTTCTCGTATCCACACCCCTTAC | 47 |

Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyValAlaGlyGluAspAlaArgSer 60
Db 474 ACCGTGTCCTCTGGAGACAAACGAGCTCTGGAGCTGGGGCCGGGGAAGAGCGCCGGTCG 533

QY 61 AspAspSerSerArgIleIleLeuGlySerAspCysAspMetHisThrGlnProTrp 80
 |||||
 Db 534 GATGACGAGCAGCGCATCATCATGATCGACTGCGATATGCACACCCAGCCGTTGG 593

Dy 81 GlnAlaIalLeuleuleuAArgProAsnGlnleueuryCyGGYalaValIeulValHisPro 100
Db 594 CAGGCCCGCTGTTCCTAAAGCCCAACCACTCTACTGGGGGGCGGTGTTGGTGCAATCCA 653

101 G1ATPheuleuhrh1a1a1a1a1sc5aaglyb1b5ValPheatgVal1a1g1e1c1y1h1s 120
 654 CAGTGGCTGCTCAGCGCCGCGCCACTGTCAGGAAGAAAGTTTTCAGAGTCCGATCTCGCCAC 713

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|
| 0y | 12 | 1y | 2y | 3y | 4y | 5y | 6y | 7y | 8y | 9y | 10y | 11y | 12y | 13y | 14y | 15y | 16y | 17y | 18y | 19y | 20y | 21y | 22y | 23y | 24y | 25y | 26y | 27y | 28y | 29y | 30y | 31y | 32y | 33y | 34y | 35y | 36y | 37y | 38y | 39y | 40y | 41y | 42y | 43y | 44y | 45y | 46y | 47y | 48y | 49y | 50y | 51y | 52y | 53y | 54y | 55y | 56y | 57y | 58y | 59y | 60y | 61y | 62y | 63y | 64y | 65y | 66y | 67y | 68y | 69y | 70y | 71y | 72y | 73y | 74y | 75y | 76y | 77y | 78y | 79y | 80y | 81y | 82y | 83y | 84y | 85y | 86y | 87y | 88y | 89y | 90y | 91y | 92y | 93y | 94y | 95y | 96y | 97y | 98y | 99y | 100y | 101y | 102y | 103y | 104y | 105y | 106y | 107y | 108y | 109y | 110y | 111y | 112y | 113y | 114y | 115y | 116y | 117y | 118y | 119y | 120y | 121y | 122y | 123y | 124y | 125y | 126y | 127y | 128y | 129y | 130y | 131y | 132y | 133y | 134y | 135y | 136y | 137y | 138y | 139y | 140y | 141y | 142y | 143y | 144y | 145y | 146y | 147y | 148y | 149y | 150y | 151y | 152y | 153y | 154y | 155y | 156y | 157y | 158y | 159y | 160y | 161y | 162y | 163y | 164y | 165y | 166y | 167y | 168y | 169y | 170y | 171y | 172y | 173y | 174y | 175y | 176y | 177y | 178y | 179y | 180y | 181y | 182y | 183y | 184y | 185y | 186y | 187y | 188y | 189y | 190y | 191y | 192y | 193y | 194y | 195y | 196y | 197y | 198y | 199y | 200y | 201y | 202y | 203y | 204y | 205y | 206y | 207y | 208y | 209y | 210y | 211y | 212y | 213y | 214y | 215y | 216y | 217y | 218y | 219y | 220y | 221y | 222y | 223y | 224y | 225y | 226y | 227y | 228y | 229y | 230y | 231y | 232y | 233y | 234y | 235y | 236y | 237y | 238y | 239y | 240y | 241y | 242y | 243y | 244y | 245y | 246y | 247y | 248y | 249y | 250y | 251y | 252y | 253y | 254y | 255y | 256y | 257y | 258y | 259y | 260y | 261y | 262y | 263y | 264y | 265y | 266y | 267y | 268y | 269y | 270y | 271y | 272y | 273y | 274y | 275y | 276y | 277y | 278y | 279y | 280y | 281y | 282y | 283y | 284y | 285y | 286y | 287y | 288y | 289y | 290y | 291y | 292y | 293y | 294y | 295y | 296y | 297y | 298y | 299y | 300y | 301y | 302y | 303y | 304y | 305y | 306y | 307y | 308y | 309y | 310y | 311y | 312y | 313y | 314y | 315y | 316y | 317y | 318y | 319y | 320y | 321y | 322y | 323y | 324y | 325y | 326y | 327y | 328y | 329y | 330y | 331y | 332y | 333y | 334y | 335y | 336y | 337y | 338y | 339y | 340y | 341y | 342y | 343y | 344y | 345y | 346y | 347y | 348y | 349y | 350y | 351y | 352y | 353y | 354y | 355y | 356y | 357y | 358y | 359y | 360y | 361y | 362y | 363y | 364y | 365y | 366y | 367y | 368y | 369y | 370y | 371y | 372y | 373y | 374y | 375y | 376y | 377y | 378y | 379y | 380y | 381y | 382y | 383y | 384y | 385y | 386y | 387y | 388y | 389y | 390y | 391y | 392y | 393y | 394y | 395y | 396y | 397y | 398y | 399y | 400y | 401y | 402y | 403y | 404y | 405y | 406y | 407y | 408y | 409y | 410y | 411y | 412y | 413y | 414y | 415y | 416y | 417y | 418y | 419y | 420y | 421y | 422y | 423y | 424y | 425y | 426y | 427y | 428y | 429y | 430y | 431y | 432y | 433y | 434y | 435y | 436y | 437y | 438y | 439y | 440y | 441y | 442y | 443y | 444y | 445y | 446y | 447y | 448y | 449y | 450y | 451y | 452y | 453y | 454y | 455y | 456y | 457y | 458y | 459y | 460y | 461y | 462y | 463y | 464y | 465 |
|----|----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|

| | | | |
|----|------|---|------|
| Db | 774 | CCCCACCTGGGCTACTCCACACCTGGGCACTTAACGACCTCATGTATCAAACCTGAAC | 833 |
| Qy | 161 | ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer | 180 |
| Db | 834 | AGAAGAATTCCTCCCACTTAAGATGTCAAGCCCATCAACGTCTCTCATATGTCCTCT | 893 |
| Qy | 181 | AlaGlyThrLysCysLeuValSerGlyArgGlyThrThrLysSerProGlnValHisPhe | 200 |
| Db | 894 | GCTGGGACAAAGTGTCTGTGTCTGTGGCTGGGGGACAAACCAAGCCCCCAAGTGCACTTC | 953 |
| Qy | 201 | ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla | 220 |
| Db | 954 | CCTAAGGTCCTCCAGTGTCTGAATATATCAAGCGGTCAAGTCAAGAAAGGTCAGAGATGCT | 1011 |
| Qy | 221 | TyrProArgGlnIleAspAspThrMetCysValAspGlyValAspLysArgAspSer | 240 |
| Db | 1014 | TACCCGAACAGATAGATGACACCACTGTCTGCGCGGAGCAAAAGCAGGTGAGACTCC | 1077 |
| Qy | 241 | CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer | 260 |
| Db | 1074 | TGCCAGGATGATTCGGGGGGGCTGTGGCTTGCAATGGCTCTCTGCAAGGACTCTGTTCTC | 1133 |
| Qy | 261 | TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys | 280 |
| Db | 1134 | TGGGGAGATTACCTTGTGCCCGGCGCCAAACAGACCGGGGTGTATACAGAACTGTGCAAG | 1197 |
| Qy | 281 | PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer | 293 |
| Db | 1194 | TTCAACCAAGTCGATCCAGAAACCACTTCAGGCAACTCC | 1232 |

RESULT 15

AX080829

| DEFINITION | Sequence 75 from Patent WO0109327. |
|------------|---|
| ACCESSION | AX080829 |
| VERSION | AX080829.1 |
| KEYWORDS | GI:13169796 |
| SOURCE | |
| ORGANISM | Homo sapiens (human) |
| | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |

REFERENCE

| FEATURES | TITLE | JOURNAL | Location/Qualifiers |
|----------|---|---------|---------------------|
| | Klajav, I. J., Laflaur, M., Mark, M. R., Masters, S. A., Pitt, R. M., Watanabe, C. K. and Wood, W. I. | | |
| | Methods of preventing the injury or death of retinal cells and treating ocular diseases | | |
| | Patent: WO 0109327-A 75 08-FEB-2001; | | |
| | Genentech, Inc. (US) | | |

```
source
1. .1570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

| ORIGIN | Alignment Scores: |
|--------|----------------------|
| | Pred. No.: 1.48e-120 |
| | Score: 1608.00 |
| | Length: 1570 |
| | Matches: 293 |

| | | |
|------------------------|---------|-----------|
| Percent Similarity: | 100.00% | Conserved |
| Best Local Similarity: | 100.00% | Mismatch |
| Query Match: | 100.00% | Indel |
| DB: | 6 | Gaps: |

US-09-936-271B-14 (1-293) x AX080829 (1-1570)

Db 354 ATGGCTACAGCAAGACCCCTGGATGGTGGGTCTGTGTCTGATCAGACGCTTGCTT 413

Gy 21 LeuG1Val1ThrG1uH1eVal1LeuH1aAsnAAspVal1SerCysAspH1sProSerAsn 40

```
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValIleGlyValAspAlaArgSer 60
DB 474 ACCGTGCTCTTGGAGACAAACGAGACTGGAGCTGGAGCCGGAGAGAGCCCGGTTCG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 534 GATGACAGCAGCAGCCGATCATCATGATCGACTGGCATATGCAACCCAGCCGTGG 593
QY 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValIleValLeuValHisPro 100
DB 594 CAGGCCCGCTGTGTGCTAAGGCCCAACAGCTTACTGCGGCGGTGTGTGTGCATCCA 653
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgIleValIleValPheArgValArgLeuGlyHis 120
DB 654 CAGTGGCTGCTCAGCGCGCCGCTGACGAGAGAGAGATTTCAGAGTCCGCTCGGCGAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 714 TACTCCCTGTACCAAGTTTATGATCTGGGACAGCATGTTCCAGGGGGTCAATCCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
DB 774 CCCCACTGCTGCTACTCCACCTGCGCACCTTAACGACCTCATGCTCAACAACTGAAC 833
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 834 AGAAGAAATGTGCTCCACTTAAGATGACAGCCCAACAGCTCTCTCATGTGCTCTCT 893
QY 181 AlIGlyThrIleCysLeuValSerGlyTyrGlyTyrThrIleSerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGTCTGTGTGTCTGGGCTGGGAGCAACCAAGGCCCCCAAGTCACTTC 953
QY 201 ProIleValIleGlnCysLeuAsnIleSerValIleuSerGlnIleValArgCysGluAspAla 220
DB 954 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGCTGCTAAGTCAGAAAAGGTGCGAGATGCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValArgAspSer 240
DB 1014 TACCGAGACAGATGATGACATCATGTCTGCGCGCGGTGACAAAGCAGGATGAGACTCC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TGGCAGGGTATTTCTGGGGGGCTGTGTGTGCAATGGCTCCCTGCAAGGACTGTGTCC 1133
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValIleTyrThrAsnLeuCysLeu 280
DB 1134 TGGGAGATTAACCTTGTGCCCCGCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1193
QY 281 PheThrIleTyrTrpIleGlnGlnIleThrIleGlnAlaAsnSer 293
DB 1194 TTCAACCAAGTGATCCAGAAACCATCCAGGCCAACTCC 1232
```

Search completed: February 26, 2005, 21:24:06
Job time : 4922 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 11:08:50 ; Search time 609 Seconds

(without alignments)
2848.086 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608

Sequence: 1 MATARPPPMVLCALITALL.....VTNLCRTKFTKIGTIOQANS 293

Scoring table:

| | |
|---------------------------|--|
| BLOSUM62 | |
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Dgapop 6.0 , Dgapext 7.0 | |
| delop 6.0 , Delext 7.0 | |

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09936271/runat.23022005.130126.27251/app_query.fasta_1.455
-DB=Geneseg.16Dec04 -OPMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09936271 @CGN_1_1_470 @runat.23022005.130126.27251 -ACPU=6 -ICPU=3
-NO_MMAP -JARGOUEY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseg.16Dec04:*

1: genesegq1980s:*

2: genesegq1990s:*

3: genesegq2000s:*

4: genesegq2001as:*

5: genesegq2001bs:*

6: genesegq2002as:*

7: genesegq2002bs:*

8: genesegq2003as:*

9: genesegq2003bs:*

10: genesegq2003cs:*

11: genesegq2003ds:*

12: genesegq2004as:*

13: genesegq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1608 | 100.0 | 882 | 2 | AAZ23318 Human PDS |
| 2 | 1608 | 100.0 | 1216 | 8 | ADA39888 Human sec |
| 3 | 1608 | 100.0 | 1216 | 10 | ADA56078 Gene enco |
| 4 | 1608 | 100.0 | 1260 | 10 | ADB80489 Ovarian c |
| 5 | 1608 | 100.0 | 1260 | 11 | ADN39197 Cancer/an |

| | | | | | | |
|----|------|-------|------|----|----------|---------------------|
| 6 | 1608 | 100.0 | 1260 | 13 | ADR25550 | Adi25550 Breast ca |
| 7 | 1608 | 100.0 | 1302 | 2 | AA57999 | Aax57999 Human BS2 |
| 8 | 1608 | 100.0 | 1302 | 2 | AA57989 | Aax57989 Human BS2 |
| 9 | 1608 | 100.0 | 1381 | 2 | AA206259 | Aax06259 Human sec |
| 10 | 1608 | 100.0 | 1381 | 10 | ADA40370 | Ada40370 Human sec |
| 11 | 1608 | 100.0 | 1381 | 10 | ADA56535 | Ada56535 Gene enco |
| 12 | 1608 | 100.0 | 1387 | 13 | ADR72632 | Adi72632 Human rev |
| 13 | 1608 | 100.0 | 1387 | 13 | ADR72874 | Adi72874 Human ova |
| 14 | 1608 | 100.0 | 1450 | 8 | ADA40367 | Ada40367 Human sec |
| 15 | 1608 | 100.0 | 1450 | 10 | ADA56532 | Ada56532 Gene enco |
| 16 | 1608 | 100.0 | 1499 | 2 | AA560578 | Aax60578 Human ker |
| 17 | 1608 | 100.0 | 1516 | 2 | AA206245 | Aax06245 Human sec |
| 18 | 1608 | 100.0 | 1516 | 8 | ADA40369 | Ada40369 Human sec |
| 19 | 1608 | 100.0 | 1516 | 10 | ADA56534 | Ada56534 Gene enco |
| 20 | 1608 | 100.0 | 1539 | 2 | AA23317 | Aaz23317 Human PDS |
| 21 | 1608 | 100.0 | 1570 | 3 | AA265070 | Aaz65070 Membrane- |
| 22 | 1608 | 100.0 | 1570 | 4 | AA521471 | Aa521471 Human CDN |
| 23 | 1608 | 100.0 | 1570 | 5 | AA44216 | Aa44216 Human PRO |
| 24 | 1608 | 100.0 | 1570 | 6 | ABK28605 | Abk28605 Human DNA |
| 25 | 1608 | 100.0 | 1570 | 8 | ACA64363 | Aca64363 Novel hum |
| 26 | 1608 | 100.0 | 1570 | 8 | ACA03830 | Aca03830 CDNA enco |
| 27 | 1608 | 100.0 | 1570 | 8 | ABX89368 | Abx89368 DNA encod |
| 28 | 1608 | 100.0 | 1570 | 8 | ABX80822 | Abx80822 Human sec |
| 29 | 1608 | 100.0 | 1570 | 8 | ACD44331 | Acdd44331 CDNA enco |
| 30 | 1608 | 100.0 | 1570 | 8 | ACD42022 | Acdd42022 Human sec |
| 31 | 1608 | 100.0 | 1570 | 8 | ABX79502 | Abx79502 Human sec |
| 32 | 1608 | 100.0 | 1570 | 8 | ACA93523 | Aca93523 Novel hum |
| 33 | 1608 | 100.0 | 1570 | 8 | ABX81205 | Abx81205 Human sec |
| 34 | 1608 | 100.0 | 1570 | 8 | ACA04251 | Aca04251 Human CDN |
| 35 | 1608 | 100.0 | 1570 | 8 | ACA93021 | Aca93021 Novel hum |
| 36 | 1608 | 100.0 | 1570 | 8 | ABX17105 | Abx17105 Human PRO |
| 37 | 1608 | 100.0 | 1570 | 9 | ACA57960 | Aca57960 Novel hum |
| 38 | 1608 | 100.0 | 1570 | 9 | ACA88409 | Aca88409 Human sec |
| 39 | 1608 | 100.0 | 1570 | 9 | ACD81916 | Acdd81916 CDNA enco |
| 40 | 1608 | 100.0 | 1570 | 9 | ADA45974 | Ada45974 Novel hum |
| 41 | 1608 | 100.0 | 1570 | 9 | ADA76405 | Ada76405 Human PRO |
| 42 | 1608 | 100.0 | 1570 | 9 | ADA19055 | Ada19055 Human PRO |
| 43 | 1608 | 100.0 | 1570 | 9 | ADA61678 | Ada61678 Homo sapl |
| 44 | 1608 | 100.0 | 1570 | 9 | ADB19463 | Adb19463 Novel hum |
| 45 | 1608 | 100.0 | 1570 | 9 | ADB28004 | Adb28004 CDNA enco |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAZ23318 | |
| ID | AAZ23318 standard; DNA; 882 BP. |
| XX | |
| AC | AAZ23318; |
| XX | |
| DT | 03-DEC-1999 (first entry) |
| XX | |
| DE | Human PDSP-1 DNA coding region. |
| XX | |
| KW | PDSP-1; Tango 114; prostate-derived serine protease; fusion protein; |
| KW | treatment; proliferative disorder; prostate cancer; screening; |
| KW | predictive medicine; growth factor; growth factor biosynthesis; |
| KW | cellular proliferation; growth factor binding protein; metastasis; |
| KW | cellular differentiation; prostate development; detection; forensic; |
| XX | serine protease; human; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO946391-A2. |
| XX | |
| PD | 16-SEP-1999. |
| XX | |
| PF | 12-MAR-1999; 99WO-US005416. |
| XX | |
| PR | 12-MAR-1998; 98US-00041400. |
| XX | |
| PA | (MILL-) MILLENNIUM BIOTHERAPEUTICS INC. |

PI Holtzman DA;
 XX
 DR WPI; 1999-561678/47.
 DR P-PSDB; AAY30524.
 XX
 PT Novel prostate derived serine protease polynucleotides and polypeptides
 PT used to modulate cellular processes.
 XX
 PS Claim 1a; Fig 1A; 94pp; English.
 XX
 CC This invention describes a novel human prostate-derived serine protease
 CC (PDSp) polynucleotide (also known as Tango 114). The prostate derived
 CC serine protease (PDSp) polypeptide is used to modulate a variety of
 CC cellular processes. It can be used to produce fusion proteins. PDSp is used
 CC to treat proliferative disorders, e.g. prostate cancer. The protein may
 CC also be used to produce antibodies, and to identify antagonists and
 CC agonists. The PDSp polynucleotides, polypeptides, homologs and antibodies
 CC can be used in screening assays; predictive medicine; and methods of
 CC treatment. PDSp cleaves growth factors, and can be used for the
 CC modulation of growth factor biosynthesis; generation of active peptides;
 CC regulation of cellular proliferation; degradation of growth factor
 CC binding proteins; regulation of cellular differentiation; regulation of
 CC metastasis; and regulation of prostate development. The PDSp
 CC polynucleotides can be used to express the protein; to detect PDSp mRNA;
 CC to detect genetic alterations in the PDSp gene; in forensic biology; and
 CC as a source of primers and probes. As serine protease have important
 CC roles in cellular processes, there exists a need for identifying novel
 CC serine protease such as the prostate derived serine protease (PDSp) of
 CC the invention. This sequence encodes the human PDSp-1 described in the
 CC invention
 XX
 SQ Sequence 882 BP; 193 A; 272 C; 237 G; 180 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7.44e-136 Length: 882
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-936-271B-14 (1-293) x AA223318 (1-882)
 QY 1 MetAlaThzAlaArgProPrtPmetTrrValLeuCyAlaLeuileThraAlaLeu 20
 Db 1 ATGGCTACAGCAAGCAACCCCTGGATGTGGTCTCTGCTGTGATCAGGCTTGCT 60
 QY 21 LeuGlyValThrgLuhIsvAlLeuAlaAsnAspValSerCyAspHisPrtSerAan 40
 Db 61 CTGGGGGTCAAGAGCATGTTCTGCGCAACATGATGTTCCCTGTGACCAACCCCTTAAC 120
 QY 41 ThrValPrtSerGlySerAenGlnAspLeuGlyValAGlyValAGlyGluAspAargSer 60
 Db 121 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAGCGCCGGCTG 180
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrgInPrtPrt 80
 Db 181 GATGACAGAGCAGCGCGCATCATCATGATCGCATTCGATATGCAACCCGCGCTGG 240
 QY 81 GlnAlaAlaLeuLeuLeuArgProAenGlnLeuTyrCySGlyValValLeuValHisPrt 100
 Db 241 CAGGCGCGCTGTGTGCTAAGGCCAACCAAGCTTACTGCGGGGGGTGGTGGCATCCA 300
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCyArgLysLysValPheArgValArgLeuGlyHis 120
 Db 301 CAGTGGCTGCTCAAGCGCGCCACTGAGAGAAAGATTTCAGAGTCCGCTCGGCCAC 360
 QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
 Db 361 TACTCCCTGTCAACAGTTTATGATCTGGGCAAGATGTTCCAGGGGGGTCAATCATC 420
 QY 141 ProHisPrtGlyTyrSerHisPrtGlyHisSerAsnAspLeuMetLeuIleLysLeuAan 160

Db 421 CCCACCTGGCTACTCCACCCCTGGCCACTGTAACGACCTCATGCTCAAACTGAAC 480
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyAsProSer 180
 Db 481 AGAAGATTGCTTCCACTTAAGATGTCAAGCCCATCAACGTCCTCTCATTTGCTCTT 540
 QY 181 AlaGlyThrLysCyLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
 Db 541 GCTGGGACAAAGTCTGTGTCTGGCTGGGGACAAACCAAGAGCCCCCAAGTGCATTC 600
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgGlyLysPala 220
 Db 601 CCTAAGTCTCTCAAGCTTGAATATCAACCGGTCTAAGTCAAAAAGGCGCAGATGCT 660
 QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspLysValArgAspSer 240
 Db 661 TACCCGAGCAATATGATATACCATTTCTTCCGCCGGGTGACAAAGCATGTAGACTTC 720
 QY 241 CySGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
 Db 721 TGCAGGGGTGATTTGGGGGGCTGTGGTGTGCAATGGCTCCCTGAGGGAGCTCGTCTC 780
 QY 261 TrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyLys 280
 Db 781 TGGGAGATTACCTTGTGCTCCGGCCCAACACACCGGGGTGTACACGAACCTTGCAAG 840
 QY 281 PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293
 Db 841 TTCACCAAGTGAATCCAGGAACATTCAGGCCAACTCC 879
 RESULT 2
 ADA39888
 ID ADA39888 standard; cDNA; 1216 BP.
 AC
 XX ADA39888;
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein encoding cDNA.
 XX
 KW Human; secreted protein; cancer; hyperproliferative disorder;
 KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
 KW anaemia; allergic reaction; asthma; cardiovascular disorder;
 KW wound healing; cyostatic; immunosuppressive; neutropic; neuroprotective;
 KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
 KW vulnerrary; cardiact; gene therapy; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO2002102993-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002WO-US008123.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 PA
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-175238/17.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating cancer or other hyperproliferative disorder,
 PT asthma, allergies or AIDS.
 XX
 PS Claim 9; SEQ ID NO 270; 3205pp; English.
 XX
 CC The invention relates to novel genes ADA39629-ADA40565 and proteins

CC ADA40566-ADA1501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1216 BP; 263 A; 388 C; 323 G; 242 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1.13e-135 | Length: | 1216 |
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 8 | Gaps: | 0 |

US-09-936-271b-14 (1-293) x ADA39888 (1-1216)

QY 1 MecLThra1aArpProTrrMetTrrVal1LeuCySa1aLeu1eThra1aLeu1e 20
DB 128 ATGGCTACAGAAACCCCGCTGATGGGTGCTGTGCTGCATCAGCAGCTTCTT 187
QY 21 Leu1yVal1Thrg1uH1eVal1eUa1aAaNa1aPVal1SerCyAaP1aPProSerAa 40
DB 188 CTGGGGGTCAAGAGCATGTTCTGCCAACAATGATGTTCTCTGTGACCACTTCAAC 247
QY 41 ThrVal1ProSerG1ySerAaP1aPProTrrMetTrrVal1LeuCySa1aLeu1e 60
DB 248 ACCGTGCTCTGAGAGCAACAGGACCTGGGAGCTGGGGGGAAGAGCCGGGCG 307
QY 61 AspAaP1aPProSerAaP1aPProTrrMetTrrVal1LeuCySa1aLeu1e 80
DB 308 GATGACAGCAAGCAACCCGATCATATGATGATGATGATGATGATGATGATGATG 367
QY 81 Gln1a1a1a1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1e 100
DB 368 CAGGCCCCGCTGTGCTAAAGGCCCAACAGCTTACGCGGGGGGGTGTGTGATCA 427
QY 101 GlnTrr1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1e 120
DB 428 CAGTGGCTGCTCAAGGCGCCGCTGAGCAAGAAAGTTTCAAGTCCGCTGGGCAC 487
QY 121 TyrSer1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1e 140
DB 488 TACTCCCTGTCACAGTTTATGAATCTGGGCAAGATGTTCCAGGGGGGTCAATTCATC 547
QY 141 ProH1aPProG1yTrrSerH1aPProG1yH1aSerAaNa1aPVal1SerCyAaP1a 160
DB 548 CCCCAACCTGGCTCTCTCCACCTGGGCACTTCAACGACCTCAATGCTCAATCACTG 607
QY 161 ArgA1g1eArpProTrr1yAaPVal1ArpPro1eAaNa1aPProSerH1aPProSer 180

DB 608 AGAAGATTGCTCCCAATAAAGATGACAGCCATCAAGCTCTCTCATTTGCCCTCT 667
QY 181 AlG1yThryScy1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1e 200
DB 668 GCTGGGCAAAAGGCTGTGTGTGCTGGCTGGGGGCAACAAAGAGCCCAAGGCACTTC 727
QY 201 Pro1yVal1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1eUa1e 220
DB 728 CTTAAGTCTCTCAAGTCTTGAATATCAAGGCTCTAAGTCTAAGAAAGTCCAGATGCT 787
QY 221 TyrPro1g1n1eAaP1aPProTrrMetPneCySa1aG1yAaP1aG1yAaP1aP 240
DB 788 TACCCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 847
QY 241 CyG1nG1yAaP1aPProTrrMetPneCySa1aG1yAaP1aG1yAaP1aP 260
DB 848 TGGCAGGCTGATCTGCGGGGCTGTGTGCTGCAATGCTCCCTGCAAGGATCTGTCTC 907
QY 261 TrrG1yAaP1aPProTrrMetPneCySa1aG1yAaP1aG1yAaP1aP 280
DB 908 TGGGAGATTAACCTGTGTGCTGCGGCCCAACAGACGGGTCTTACACGAACCTGTGCAAG 967
QY 281 PheThryTrrP1eG1nG1yThry1eG1nG1yThry1eG1nG1yThry1eG1nG1y 293
DB 968 TTCACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCC 1006

RESULT 3

ADA56078 standard; DNA; 1216 BP.

ADA56078;

20-NOV-2003 (first entry)

Gene encoding human secreted protein #257.

KW immunosuppressive; antiinflammatory; antisthmatic; antiallergic;
KW cytoprotective; cerebroprotective; neuroprotective; neurotropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology; ds;
KW gene.

XX Homo sapiens.

XX WO2002102994-A2.

XX 27-DEC-2002.

XX 19-MAR-2002; 2002WO-US008278.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-167512/16.

XX P-PSDB; ADA56974.

PT New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.

XX Claim 21; SEQ ID NO 267; 1754bp; English.

XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that binds to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to a gene encoding one of the polypeptide of the invention. Note: The
CC sequence data for this patent did form part of the printed specification,
CC but was obtained in electronic format directly from WIPD at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1216 BP; 263 A; 388 C; 323 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Score: 1.13e-135 Length: 1216
Percent Similarity: 1608.00 Matches: 293
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: 10 Indels: 0 Gaps: 0

US-09-936-271B-14 (1-293) x ADA56078 (1-1216)

QY 1 MetAlaThrAlaArgProPrtPmetTTPValLeuCyAlaLeuIleThrAlaLeu 20
DB 128 ATGGCTACAGCAAGACCCCTGGATGGGTCTCTGCTGTATCACAGCTTGCTT 187
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSer 40
DB 188 CTGGGGGTACACAGCATGTTCTGGCCAACAATGATGTTCCGTGACCACTCTTAAC 247
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValAGlyValGluAspAlaArgSer 60
DB 248 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAGCGCCGCTG 307
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTP 80
DB 308 GATGACAGAGCAGCGGCATCATCATGATCGACATGATGACACCCGCGCTGG 367
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyTCysGlyAlaValLeuValHisPro 100
DB 368 CAGGCGCGCTGTGTAAAGGCCAACACAGCTCTACCTGCGGGCGGTGTGTGATCA 427
QY 101 GlnTPLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
DB 428 CAGTGGCTGTCTACGGCGCGCCACTGACGAGAAAGAGTTTCAGAGTCCGTCTCGGCCAC 487
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValTyrSerIle 140
DB 488 TACTCCCTGTACCACTTATGATCTGGGACGACATGTTCCAGGGGGGTCAAATCATC 547
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleTyrLeuAsn 160

DB 548 CCCACCTGTGCTACTCCACCCCTGGGCACTCACTTAACGACTCATGCTCAACAACTGAAC 607
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 608 AGAAGATTCGTCCCACTAAAGATGCAGACCCATCAAGTCTCTCATTTGCTCTCT 667
QY 181 AlaGlyThrIleCysLeuValSerGlyTTPGlyThrThrIleSerProGlnValHisPhe 200
DB 668 GCTGGACAAAGTCTGTGCTGTGCTGGGGGACACCAAGACCCCAAGTCACTTC 727
QY 201 ProValLeuGlnCysLeuAsnIleSerValIleSerGlnValArgCysGluAspAla 220
DB 728 CTTAAGTCTCTCAAGTCTTGAATATCAGCGTCTTAAGTCAAGAAAGGTGCGAGATGCT 787
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspValAlaGlyArgAspSer 240
DB 788 TACCCGACAGATATGATATACACCATGTTCTGCGCGGACAAAGCAGTGAAGCTCC 847
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 848 TCGCAGGATGATTTGGGGGGCTGTGTGCTGCATGCTCCCTGAGGGACTCGTGTCC 907
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIys 280
DB 908 TCGGGGATTAACCTTGTGCTCCGCGCCAAACAGACGGGGTGTACACGAACCTTGCAAG 967
QY 281 PheThrIleTyrIleGlnIleGluThrIleGlnIleAsnSer 293
DB 968 TTCACCAAGTGAATCCAGGAACCATCCAGGCCAACTCC 1006

RESULT 4
ADB80489
ID ADB80489 standard; DNA; 1260 BP.
XX AC ADB80489;
XX DT 04-DEC-2003 (first entry)
XX DE Ovarian cancer-associated transcript #27.
XX KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
XX KW post-operative chemotherapy; radiation therapy; tumour prognosis;
XX KW pre-cancerous lesion detection; de; gene.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 43..924
XX FT /*tag= a
XX PN MO2002102235-A2.
XX PD 27-DEC-2002.
XX PF 18-JUN-2002; 2002WO-US019297.
XX PR 18-JUN-2001; 2001US-0299234P.
XX PR 27-AUG-2001; 2001US-0315287P.
XX PR 05-SEP-2001; 2001US-0317544P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (E05B-) EOS BIOTECHNOLOGY INC.
XX PI Mack DH, Gish KC;
XX DR WPI; 2003-167431/16.
XX DR P-PSDB; ADB80490.
XX PT Detecting an ovarian cancer-associated transcript in a cell from a
XX PT patient; comprises contacting a biological sample from the patient with a
XX PT polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 10; Page 292; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-associated transcript in a cell from a patient, by contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to any of one of 80 nucleic acid sequences given in the specification. The method is useful in diagnosing ovarian cancer and in identifying and using agents and/or targets that inhibit ovarian cancer. The nucleic acid molecule, polypeptide, and the antibody may also be used in detecting ovarian cancers, monitoring and early detection of relapse following treatment, monitoring response to therapy, selecting patients for post-operative chemotherapy or radiation therapy, in selecting mode of therapy, determining tumour prognosis, early detection of pre-cancerous lesions, and as vaccines. This sequence corresponds to one of the nucleic acids used for the detection method of the invention.

SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.18e-135 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-936-271b-14 (1-293) x ADB80489 (1-1260)

QY 1 MeAlaThzAlaArgProPOTrMetTrpValLeuCyAlaLeuIleThAlaLeu 20
Db 43 ATGGCTACAGACAGACCCCTGTGATGGGTCTGTCTGTGATCAGACCTTCT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 103 CTGGGGGTCAAGAGCATGTTCTGCCAACAATATGTTCTGTGACCACTTCAAC 162
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValAlaGlyValAspAlaArgSer 60
Db 163 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGGAGAGAGCGCGGTGG 222
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 223 GATGACGACGACGCCCATCATCAATGATCCAGTCCGATATGACACCCGCTGG 282
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100
Db 283 CAGGCGCGCTGTGTGCTAAAGGCCCAACAGCTTACTGGGGGGGGTGTGTGCATCA 342
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
Db 343 CAGTGGGTGTCAAGCGCGCCCACTGACAGAAAGAAATTTTCAAGATCCGTCTGGGCAC 402
QY 121 TyrSerLeuSerProValTrpGluSerGlyGlnGlnMetPheGlnGlyValValSerTle 140
Db 403 TACTCCCTGTCAACAGTTTATGATCTGGGCAAGAGATTTCCAGGGGGGTCAATTCATC 462
QY 141 ProHisProGlyTrpSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
Db 463 CCCCACTGGTACTCCCACTGGCCACTCTAAACAACCTCATGCTCAACAACTGAAC 522
QY 161 ArgArgIleArgProTrpTrpValArgProIleAsnValSerSerHisCysProSer 180
Db 523 AGAAGAAATGTCTCCACTAAAGATGACAGACCCATCAACGCTCTCTCATGTGCTCTCT 582
QY 181 AlaGlyThrIleCysLeuValSerGlyTrpGlyThrThrIleSerProGlnValHisPhe 200
Db 583 GCTGGGCAAGAGTCTGTGTGTGTGGTGGGGGCAACCAAGAGCCCCCAAGTGCATCTC 642
QY 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
Db 643 CTTAAGGTCTCTCCAGTCTTGAATATCAAGCGTGTAAAGTCAAGAAAGTGGAGAGTCT 702

QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValArgAspSer 240
Db 703 TACCCAGACAGATGATATGACACCATCTTCTGGCGCGGTGACAAAGCATAGACTCC 762
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnLeuValSer 260
Db 763 TGCACAGGTGATCTTGGGGGGGCTGTGTGTGTCGCAATGGCTCCCTGCAGGGAATCGTTC 822
QY 261 TrpGlyAspTrpCysValAlaArgProAsnArgProGlyValTrpThrAsnLeuCys 280
Db 823 TGGGAGATTAACCTGTGTGCGCCCAACAGACCGGGGTCTTACAGAACTTGTCAAG 882
QY 281 PheThrIleTrpIleGlnGluTrpIleGlnAlaAsnSer 293
Db 883 TTCAACAGTGTATCCAGAAACATTCAGGCCAATCTCC 921

RESULT 5
ADN39197
ID ADN39197 standard; cdna, 1260 BP.

AC ADN39197;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.

XX Human, differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiact; immunomodulatory;
vulnerable; gene therapy; vaccine; gene; ss.

OS Homo sapiens.
XX
PN WO2003042661-A2.

XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.

XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386514P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397755P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.

XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevez PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI; 2003-468649/44.
XX P-PSDB; ADN39198.

XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.

PS Claim 8; SEQ ID NO 515; 1385bp; English.

XX The invention relates to nucleic acids and proteins (ADN3683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularization syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

XX
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-135 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271B-14 (1-293) x ADN39197 (1-1260)

QY 1 MetAlaThrAlaArgProProTpmcTTPvAlleuCyAlaLeuIleThrAlaLeu 20
DB 43 ATGGCTACAGACAGACCCCTGATGTGGTCTGTGCTGTGATCACAGCTTGCT 102
QY 21 LeuGlyValThrGluHisValleuAlaAsnAspValSerCyAspHisProSerAa 40
DB 103 CTGGGGGTACAGAGATGTTCTGCCCAATGATGTTCTGTACCACTCTTAAAC 162
QY 41 ThrValProSerGlySerArgGlnAspLeuGlyAlaGlyValAspAlaAspSer 60
DB 163 ACCGGCCCTCTGGGAGCAACGAGACCTGGGAGCTGGGAGCCGGGAAAGCCGCCGTG 222
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTP 80
DB 223 GATGACAGACAGCCGCGCATCATCATGATCGACTGCGGATATGCAACCCGCGTGG 282
QY 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValAlaValleuValHisPro 100
DB 283 CAGGCGCGCTGTGTAAAGGCCCAACAGCTTACTGCGGGCGGTGGTGGATCCA 342
QY 101 GlnTPLeuLeuThrAlaHisGlyValGlyValPheArgValArgLeuGlyHis 120
DB 343 CAGTGGCTCTACAGCGCCGCCACTGACAGAAAGATTTCAGATCCGTCTCGGCCAC 402
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValHisSerIle 140
DB 403 TACTCCCTGTACCACTTTATGATCTGGGACAGCAATGTTCCAGGGGGTCAATATCATC 462
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleHisLeuAsn 160
DB 463 CCCACCTCTGGCTACTCCACCTGGCCACTTAAGAGACTTATGCTCATCAACAGAAC 522
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyAspProSer 180
DB 523 AGAAGATTCTGCCCTAAAGATGTCAGACCCATCAAGTCTCTCATCTCTCCCTCT 582
QY 181 AlaGlyThrLeuCyAlaValSerGlyTTPGlyThrThrHisSerProGlnValHisPhe 200
DB 583 GCTGGAGACAAAGTCTGTGTGTGGCTGGGAGCAACCAAGAGCCCTCAAGTGCATCTTC 642
QY 201 ProLysValleuGlnCysLeuAsnIleSerValleuSerGlnLysArgCyGluAspAla 220

DB 643 CCTAAGTCTCTCAGAGCTTGAATATACAGCGTGTAAAGTCAAGAAAGGTCAGAGATGCT 702
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValGlyArgAspSer 240
DB 703 TACCCGAGCAATATGATATACACATGTTCTGCGCGGAGCAAAAGCAGTGAAGCTTC 762
QY 241 CySGlnGlyAspSerGlyValProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 763 TCCAGAGGATTTCTGGGGGCTGTGTGTGCTGCAATGCTCTCGAGGAGACTCGTGTCC 822
QY 261 TTPGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyHis 280
DB 823 TGGGAGATTACCTTGTGCTGCGGCCCAACAGACCGGGGTGTACACGAACCTTGCAAG 882
QY 281 PheThrIleTyrIleGlnGluThrIleGlnAlaAsnSer 293
DB 883 TTCACCAAGTGTATTCAGAAACCATCCAGGCCAATCTCC 921

RESULT 6

ADR25550
ID ADR25550 strand; DNA; 1260 BP.

XX AC ADR25550;
XX 21-OCT-2004 (first entry)
XX
XX DE Breast cancer prognosis marker #1411.
XX
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX
XX Homo sapiens.

OS
XX
XX PN WO2004065545-A2.
XX
XX PD 05-AUG-2004.
XX
XX PF 15-JAN-2004; 2004WO-US001100.
XX
XX PR 15-JAN-2003; 2003US-00342887.

XX
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX (NECA-) NETHERLANDS CANCER INST.
XX
XX PI Van't Veer LJ, He Y;
XX
XX DR WPI; 2004-593473/57.

XX
XX PT Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX
XX
XX PS Disclosure, SEQ ID NO 1411; 226bp; English.

XX
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.

XX
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-135 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

| ID | DB | 13 | Gaps: | 0 |
|---|-----|--|-------|---|
| US-09-936-271B-14 (1-293) x ADR25550 (1-1260) | | | | |
| QY | 1 | MechAlaThrAlaArgProPOTIPMeCTrPValLeuCyAlaLeuIleThrAlaLeuLeu | 20 | |
| Db | 43 | ATGGCTACACAGAACCCCTGGATGTGGGTCTCTGTGCTCTGATCAGACCTGCTT | 10 | |
| QY | 21 | LeuGIyValIThrGIuHISValLeuAlaAmaAmaPValSerCyAspHisProSerAn | 40 | |
| Db | 103 | CTGGGGGTCAcAGAGCATGTTCTCGCCAAcAATGATGTTTCTGTGACCAcCCCTCTAC | 16 | |
| QY | 41 | ThrValProSerGIySerAenGIyAmaPLeuGIyAlaGIyAlaGIyLysAlaArgSer | 60 | |
| Db | 163 | ACGGTGGCTCTGGAGCAACAGACCTGGAGCTGGGGCGGGAGAGACGCCGGGTGG | 22 | |
| QY | 61 | AspAspSerSerSerArgIleIleAenGIySerAspCyAspMetHisThrGIyProTrp | 80 | |
| Db | 223 | GATACACGACGACGACCGCATCATCATGATCCGACTGGAGTATGCAcCCACCGCTGG | 28 | |
| QY | 81 | GIuAlaAlaLeuLeuLeuArgProAenGIyLeuTYrCyGIyAlaValLeuValHisPro | 100 | |
| Db | 283 | CAGGCCCGCTGTGCTTAAGGCCCAACAGCTCTACTGGGGGGCGGTGTGGTGAATCCA | 34 | |
| QY | 101 | GIuITrPLeuLeuThrAlaAlaHisGyAlaGlyblyValPheAlaTyValArgLeuGIyHis | 120 | |
| Db | 343 | CAGTGGGTGTCTCACGGCCGCCCTGCACAGAAAGAAAGTTTTCAGAGTCCGTCTGGCTAC | 40 | |
| QY | 121 | TyrSerLeuSerProValTYrGIuSerGIyGIuGIyMetPheGIuGIyValIySerIle | 140 | |
| Db | 403 | TACTCCCTGTCAcCATGTTTATGATTTGGGACGCAAGATGTTTCAAGGGGTCAATTCATC | 46 | |
| QY | 141 | ProHisProGIyTYrSerHisProGIyHisSerAmaAspLeuMetLeuIleValLeuAsn | 160 | |
| Db | 463 | CCCCACCCCTGGCTACTCCACCCCTGGCCACTCTAACAGACCTCATGCTCAAACTGAC | 52 | |
| QY | 161 | ArgArgIleArgProThrIyAspValAlaArgProIleAenValSerSerHisCyProSer | 180 | |
| Db | 523 | AGAAGAAATTCGTCCCACTAAAGTGCAGACCCATCAcGTCTCTCATTTGTCCCTCT | 58 | |
| QY | 181 | AlAGIyThrIyAsCyAlaLeuValSerGIyTYrPGLyThrIySerProGIyValHisPhe | 200 | |
| Db | 583 | GCTGGACAAAGTCTTGTTGTGTGTGTGGGGGACAAcCAAGGCCCCAAAGTCACTTC | 64 | |
| QY | 201 | ProIyValIleuGIyCyAlaLeuAsnIleSerValLeuSerGIyIyAsyCyGIyAspAla | 220 | |
| Db | 643 | CTTAAGGCTCTCAAGTGTGAATATCAcGGTGTCAAGAAAGGTGGAGAGTCT | 70 | |
| QY | 221 | TyrProArgGIyIleAmaPThrMetPheCyAlaGIyAspIyValGIyAlaArgAspSer | 240 | |
| Db | 703 | TACCCGACAGATATATATACCAcAATGTTCTGCGCGGTGCACAAAGCAGTATGAGACTCC | 76 | |
| QY | 241 | CyAGIuGIyAspSerGIyGIyProValValCyAaenGIySerLeuGIyIyLeuValSer | 260 | |
| Db | 763 | TGCCAGGTGATTTCTGGGGGGCTGTGGTCTGCATATGCTCCCTGCAGAGGACTGTCTC | 82 | |
| QY | 261 | TrpGIyAspTYrProCyAlaArgProAmaArgProGIyValTYrThrAsnLeuCyAla | 280 | |
| Db | 823 | TGGGAGATTACCTTGTGTGCGCGCCCAACACACCGGGTGTCTACACAACTCTGCAG | 88 | |
| QY | 281 | PheThrIyTrpIleGIuGIyThrIleGIuAlaAenSer | 293 | |
| Db | 883 | TTCAcCAAGTGTATCCAGGAAACCATCCAGGCCAATCTCC | 921 | |
| RESULT 7 | | | | |
| AAK57990 standard; DNA; 1302 BP. | | | | |
| AAK57990; | | | | |
| AAK57990; | | | | |
| 19-JUL-1999 (first entry) | | | | |
| Human BS247 specific polynucleotide #8. | | | | |

| | |
|---|---|
| XX | BS247; detection; diagnosis; breast cancer; atypical hyperplasia; |
| KM | fibroadenoma; cystic breast disease; gene therapy; ss. |
| XX | |
| OS | Homo sapiens. |
| PN | M09922027-A1. |
| XX | |
| PD | 06-MAY-1999. |
| XX | |
| PF | 28-OCT-1998; 98MO-US022906. |
| XX | |
| PR | 28-OCT-1997; 97US-0063431P. |
| PR | 28-OCT-1997; 97US-00968838. |
| XX | |
| PA | (ABBOTT) ABBOTT LAB. |
| XX | |
| PI | Billing-Medel PA, Cohen M, Colpitts TL, Friedman EN, Gordon J; |
| P1 | Grenados SN, Hodges SC, Klass MR, Kratochvil JD, Russell JC; |
| PI | Strope DN, Yu H; |
| XX | |
| DR | WPI, 1999-312977/26. |
| XX | |
| PT | Breast tissue derived cDNA contig and consensus polypeptide sequence. |
| PS | |
| XX | Claim 1; Page 106; 112pp; English. |
| CC | This sequence is a BS247 specific polynucleotide. The invention relates |
| CC | to a method of detecting the presence of a target BS247 polynucleotide, |
| CC | especially mRNA, in a test sample. BS247 polynucleotides are derived from |
| CC | breast tissue. The polynucleotides, polypeptides or antibodies are useful |
| CC | for providing information leading to the detection, diagnosis, staging, |
| CC | monitoring, prognosis, in vivo imaging, prevention or treatment, |
| CC | determining predisposition to, diseases and conditions of the breast, |
| CC | such as breast cancer, atypical hyperplasia, fibroadenoma and cystic |
| CC | breast disease. Drug treatment or gene therapy for breast cancer, can be |
| CC | based on these identified gene sequences and the efficacy of any |
| CC | particular therapy can be monitored. The BS247-derived reagents are |
| CC | advantageous for detection of breast cancer due to their specificity. The |
| CC | reagents also provide an alternative, non-surgical diagnostic method |
| CC | capable of detecting early stage breast disease, such as cancer |
| SQ | Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other; |
| Alignment Scores: | |
| Pred. No.: | 1, 23e-135 Length: 1302 |
| Score: | 1608.00 Matches: 293 |
| Percent Similarity: | 100.00% Conservative: 0 |
| Best Local Similarity: | 100.00% Mismatches: 0 |
| Query Match: | 100.00% Indels: 0 |
| DB: | 2 Gaps: 0 |
| US-09-936-271B-14 (1-293) * MAX57990 (1-1302) | |
| DY | 1 MetAlaThrAlaArgProFroTrpMetTTPValLeuCyAlaLeuIleThrAlaLeu 20 |
| Db | 88 ATGGCTACAGCAAGACCCTCGATGTGGGTCCTGCTCATGCATCAGCCTTGCTT 14 |
| DY | 21 LeuGIyValThGlUHisValLeuAlaAsnAsnAPVAlSerCyAaPhIPProSerAsn 40 |
| Db | 148 CTGGGGGGTACAGAGAGATTCTCCCAACAATGATGTTCCTGTGACACCCCTTTAAC 207 |
| DY | 41 ThrValProSeGIySerAnGlnInAPreUGlyAlaGlyAlaGlyValaSPAlaArgSer 60 |
| Db | 208 ACCGTGCCCTCTGGAGACCAACAGACTGGGAGCTGGGAGCCGGGGAAGAAGCCGGTGG 267 |
| DY | 61 AspASPserSerSerArgIlelleAnGlySerAPCyAaPMechIstHngInProTrp 80 |
| Db | 268 GATGACAGCAGCAGCGCGATCATCAATGATTCGACTGCGCATGTCACACCCAGCGCTGG 327 |
| DY | 81 GlnAlaAlaLeuLeuLeuArgProAnGlnLeuTyrcYagIyAlaValLeuValHisPro 100 |
| Db | 328 CAGGCCGCGCTGTGTGTAAAGCCCAACAGCTACTGCGGGGGGGTGTGTGCATCCA 387 |

```
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyLeuValPheArgValArgLeuGlyHis 120
XX |||
PS 388 CAGTGGCTCTCAGCGCCGCCCTCAGAGAAAGATTTCAGATCCGCTCGGCCAC 447
DB |||
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
XX |||
PS 448 TACTCCCTGTACACAGTTTATGAAATCTGGGCGACAGATGTTCCAGAGGGGCTCAATCCATC 507
DB |||
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLysLeuAn 160
XX |||
PS 508 CCCACCCCTGGCTACTCCACCCCTGGCCACTTAACGACCTCAATGCTCAATCAACGTAAC 567
DB |||
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
DB 568 AGAAGAAATTCGTCCCTCAATAGATGTCAGACCATCACTCTCTCTATGTCCTCTCT 627
QY 181 AlaGlyThrIlyCyLeuValSerGlyTyrPglYThrThrIlySerProGlnValHisPhe 200
DB 628 GCTGGGACAAGGCTTGGTGTCTGGCTGGGGGACAAACAGAGCCCCCAAGTGCACTTC 687
QY 201 ProIlyValLeuGlnIlyCyLeuAnIleSerValLeuSerGlnIlyArgCysGlyAspAla 220
DB 688 CTTAAGTCTCTCAGTGTGCTTGAATATCAGCGTCTAAGTCAAGAAAGGTGCGAGATGCT 747
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIlyAspIlyArgAspSer 240
DB 748 TACCCAGACAGATATGATGACACCATGTTCTGGCCGGGTGACAAACAGGTAGAGACTCC 807
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAnGlySerLeuGlnGlyLeuValSer 260
DB 808 TGCACAGGTGATCTGGGGGGCTGTGTGCTGCAATGGCTCCCTGCAGGGGACTCGTGTCC 867
QY 261 TrpGlyAspThrProCysAlaArgProAnArgProGlyValTyrThrAnLeuCysLys 280
DB 868 TGGGAGATTAACCTTGTCGCCGCCCAACAGACCGGGGTCTACAGAACTCTCAAG 927
QY 281 PheThrIlySerIleGlnIlyThrIleGlnIleAlaAnSer 293
DB 928 TTCACCAAGTGAATCCAGGAAACATCCAGGCCCAACTCC 966
RESULT 8
AAK57989 standard; DNA; 1302 BP.
XX
AC AAK57989;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human BS247 specific polynucleotide #7.
XX
KM BS247: detection; diagnosis; breast cancer; atypical hyperplasia;
KM fibroadenoma; cystic breast disease; gene therapy; ss.
XX
OS Homo sapiens,
XX
PN WO922027-A1.
XX
PD 06-MAY-1999.
XX
PF 28-OCT-1998; 98WO-US022906.
XX
PR 28-OCT-1997; 97US-0063431P.
XX
PR 28-OCT-1997; 97US-00968836.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Grandos EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC,
PI Stroepe SD, Yu H;
XX
DR WPI; 1999-312977/26.
XX
```

```
PT Breast tissue derived cDNA contig and consensus polypeptide sequence.
XX
PS Claim 1, Page 105, 112pp: English.
XX
CC This sequence is a BS247 specific polynucleotide. The invention relates
CC to a method of detecting the presence of a target BS247 polynucleotide,
CC especially mRNA, in a test sample. BS247 polynucleotides are derived from
CC breast tissue. The polynucleotides, polypeptides or antibodies are useful
CC for providing information leading to the detection, diagnosis, staging,
CC monitoring, prognosis, in vivo imaging, prevention or treatment,
CC determining predisposition to, diseases and conditions of the breast,
CC such as breast cancer, atypical hyperplasia, fibroadenoma and cystic
CC breast disease. Drug treatment or gene therapy for breast cancer, can be
CC based on these identified gene sequences and the efficacy of any
CC particular therapy can be monitored. The BS247-derived reagents are
CC advantageous for detection of breast cancer due to their specificity. The
CC reagents also provide an alternative, non-surgical diagnostic method
CC capable of detecting early stage breast disease, such as cancer
XX
SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,23e-135 Length: 1302
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-936-271B-14 (1-293) x AAK57989 (1-1302)
QY 1 MetAlaThrAlaArgProProThrMetTrpValLeuCyAlaLeuIleThrAlaLeuLeu 20
DB 88 ATGGCTACAGCAAGACCCCTCGATGGGGTCTGTGCTGTATACACACCTTGCT 147
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAn 40
DB 148 CTGGGGGTTCACAGAGCATGTTCTCCCAACATGATGTTTCTGTGACACCCCTCTAC 207
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSer 60
DB 208 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGCGGGAGAGACCCCGATCG 267
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 268 GATGACAGAGCAGCGCATCATCATGATGATCGATGCAATCAACCCAGCGCTGG 327
QY 81 GlnAlaAlaLeuLeuLeuArgProAnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
DB 328 CAGGCCGCGCTGTGCTTAAGGCCCAACAGCTCTACTGGCGGGCGGTGTGGTCATCCA 387
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyLeuValPheArgValArgLeuGlyHis 120
DB 388 CAGTGGCTCTCAGCGCCGCCCTCAGAGAAAGATTTCAGATCCGCTCGGCCAC 447
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
DB 448 TACTCCCTGTACACAGTTTATGAAATCTGGGCGACAGATGTTCCAGAGGGGCTCAATCCATC 507
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLysLeuAn 160
DB 508 CCCACCCCTGGCTACTCCACCCCTGGCCACTTAACGACCTCAATGCTCAATCAACGTAAC 567
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
DB 568 AGAAGAAATTCGTCCCTCAATAGATGTCAGACCATCACTCTCTCTATGTCCTCTCT 627
QY 181 AlaGlyThrIlyCyLeuValSerGlyTyrPglYThrThrIlySerProGlnValHisPhe 200
DB 628 GCTGGGACAAGGCTTGGTGTCTGGCTGGGGGACAAACAGAGCCCCCAAGTGCACTTC 687
QY 201 ProIlyValLeuGlnIlyCyLeuAnIleSerValLeuSerGlnIlyArgCysGlyAspAla 220
XX |||
```


Db 969 TTCCACCAAGTGATTCGAGAACCATTCACGGCCAACTCC 1007

RESULT 10
ID ADA40370 standard; cDNA; 1381 BP.
XX ADA40370;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein encoding cDNA.
XX
KW Human; secreted protein; cancer; hyperproliferative disorder;
KM rheumatoid arthritis; autoimmune disorder; hematopoietic disorder;
KM anaemia; allergic reaction; asthma; cardiovascular disorder;
KM wound healing; cyclostatic; immunosuppressive; nootropic; neuroprotective;
KM antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
KM vulnerrary; cardiant; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO2002102993-A2.
XX
PD 27-DEC-2002.
XX
PF 19-MAR-2002; 2002WO-US008123.
XX
PR 21-MAR-2001; 2001US-0277340P.
XX PR 19-JUL-2001; 2001US-0306171P.
XX PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HDMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-175238/17.
XX
PT New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.
XX
PS Claim 9; SEQ ID NO 752; 3205bp; English.
XX
XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunochemical probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published pct sequences.

[illegible]

RESULT 11
ADA56535
ID ADA56535 standard; DNA; 1381 BP.
XX
XX ADA56535;
XX
DT 20-NOV-2003 (first entry)
XX
XX Gene encoding human secreted protein #257.
DE
XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
XX cytoskeletal; cerebroprotective; neuroprotective; neurotropic;
XX cardiovascular; antiatherosclerotic; gene therapy;
XX human secreted protein; immune disorder; inflammation;
XX respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
XX inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
XX multiple sclerosis; ischemic brain injury; Parkinson's disease;
XX Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
XX triple helix formation; antisense gene therapy; forensic biology; ds;
XX gene.
XX
XX Homo sapiens.
OS
XX
XX MO2002102994-A2.
PN
XX
XX 27-DEC-2002.
PD
XX
XX 19-MAR-2002; 2002MO-US0082378.
PF
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI; 2003-167512/16.
XX P-PSDB; ADA57428.
DR
XX
XX New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX
XX
XX Claim 21; SEQ ID NO 724; 1754bp; English.
PS
XX
XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that bind to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorder, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to a gene encoding one of the polypeptide of the invention. Note: The

CC sequence data for this patent did form part of the printed specification,
CC but was obtained in electronic format directly from WIPD at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,336-135 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-936-271B-14 (1-293) x ADA56535 (1-1381)
QY 1 MetAlaThrAlaArgProPOTRMETTPValLeuCYsAlaLeuIleThrAlaLeu 20
DB 129 ATGGCTACAGCAAGACCCCTGGATGGGTCTGTCTCTGATCACAGCTTGCTT 188
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 189 CTGGGGGTCAAGAGCATGTTCTCGCCAAATGATGTTCTGTGACCACTCCCTTAC 248
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValAspAlaArgSer 60
DB 249 ACCGTGCTCTTGGAGCAAGACCTGGAGCTGGGGCGGAGAGCGCCGGTTCG 308
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 309 GATGACAGCAGCAGCGGCATCATCATGATCCGACGTGATGACACCAAGCGCTGG 368
QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB 369 CAGGCCCGCTGTGTGTAAGCCCAACCACTCTACTCGGGCGGTGTGTGCATCCA 428
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyLeuValPheArgValArgLeuGlyHis 120
DB 429 CAGTGGCTGTCAAGCGGCCCACTGCGAGAAAGAACTTTTCAGAGTCCGTTCGGCAC 488
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValHisSerIle 140
DB 489 TACTCCCTGTACACAGTTTATGATCTGGGCACAGATGTTCCAGGGGGTCAATCATC 548
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuIleValLeuAsn 160
DB 549 CCCCACTGTGTACTCCCACTGGCCACTTACACCTCATGCTCATCAAACTCAAC 608
QY 161 ArgArgIleArgProThrHisAspValArgProIleAsnValSerSerHisCyProSer 180
DB 609 AGAAGATTGCTCCCACTAAAGATGACCACTCAACCTCTCTCTCTCTCTCTCT 668
QY 181 AlaGlyThrIleCySerLeuValSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
DB 669 GCTGGGACAAAGTCTTGCTGTCTGGTGGGACAAACAAAGCCCCCAAGGCACTTC 728
QY 201 PolyValLeuGlnCySerLeuAsnIleSerValLeuSerGlnIleAspValAspAla 220
DB 729 CTTAAGGTCTCTCAAGTCTTGAATATCAAGCTGTATGATGAGAAAGTCCAGGAGCTT 788
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAspValAspSer 240
DB 789 TACCCGAGACAGATGATGACACCATGTTCTGGCCGGTATCAAAAGAGTGAAGATCC 848
QY 241 CysGlnIleAspSerGlyGlyProValValCyAsnGlySerLeuGlnIleValAspSer 260
DB 849 TGCAGGATATTTGGGGGGCTGTGTGTGCAATGCTCCCTGCAAGGAGACTGTCTCC 908
QY 261 TrpGlyAspTyrProCyValAlaArgProAsnAspProGlyValIleThrAsnLeuCyVal 280
DB 909 TGGGAGATTAACCTTGTGCGGGCCCAACAGACCGGGTGTCTTACAGAACTTCTGCAAG 968
QY 281 PheThrIleTrpIleGlnGluThrIleGlnAlaAsnSer 293

XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 1.
 DE kallikrein 5, tumour marker; ovarian cancer;
 KM epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
 KW KLK5; ds; gene.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 170..1051
 FT /*tag= a
 FT /product= "Human ovarian cancer-related tumour marker
 FT kallikrein 5 (hks) protein"
 FT
 PN MO2004075713.A2.
 XX
 PD 10-SBP-2004;
 XX
 PF 26-FEB-2004; 2004WO-CA000281.
 XX
 PR 26-FEB-2003; 2003US-0450406P.
 XX
 PA (MOUN) MOUNT SINAI HOSPITAL.
 XX
 PI Diamandis EP.
 XX
 DR WPI; 2004-661815/64.
 DR P-PSDB; ADR72873.
 XX
 PT Kallikrein markers detection method for detecting ovarian cancer in
 PT patient, involves detecting kallikrein markers and CA125 in sample
 PT obtained from patient, and comparing detected amounts with standard
 PT amounts.
 XX
 PS Example 2; SEQ ID NO 4; 102bp; English.
 XX
 CC The invention relates to a novel method for detecting a plurality of
 CC kallikrein markers associated with ovarian cancer. The method comprises
 CC obtaining a sample from a patient and detecting in the sample a plurality
 CC of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
 CC the kallikrein markers are selected from the group consisting of
 CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
 CC kallikrein 11. The detected amounts of the kallikrein markers are
 CC compared with standard amounts. The method of the invention may be useful
 CC for detecting kallikrein markers associated with ovarian cancer in a
 CC patient and thus for detecting ovarian cancer, particularly epithelial
 CC ovarian carcinoma. The current sequence is that of the human ovarian
 CC cancer-related tumour marker kallikrein 5 (KLK5) DNA 1 of the invention
 CC which encodes a secreted serine protease and is located at chromosome
 CC 19q13.4.
 CC
 SQ Sequence 1387 BP; 290 A; 437 C; 351 G; 309 T; 0 U; 0 Other;
 XX
 XX
 Alignment Scores:
 Pred. No.: 1,33e-135 Length: 1387
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-936-271B-14 (1-293) x ADR72874 (1-1387)
 QY 1 MetAlaThrAlaArgProTyrMetTyrValIleuCYsAlaIleuIleThrAlaIleu 20
 DB 170 ATGGCTACACAGAACCCCTGGATGTGGTGTCTGTGTGTGATCACAGCTTCTT 229
 QY 21 LeuGLyValThrGLuHISValIleuAlaAsnAspValSerCYsAspHisProSerAsn 40
 DB 230 CTGGGGGTCAACAGGAGTGTCTGCCAACATGATATTTCTCTGTGACCAACCCCTTAAC 289
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyAlaGlyIuAspAlaArgSer 60

DB 290 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAGACGCCGCTCG 349
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCYsAspMetHisThrGlnProTyr 80
 DB 350 GATGACAGACAGACCGCATCATCAATGATCGAGTGTGACATGACACCCAGCGGTGG 409
 QY 81 GlnAlaAlaIleuIleuAspProAsnGlnIleuTyrCYsGlyValAlaValIleuValHisPro 100
 DB 410 CAGCGCGCTGTGTGCTTAAGGCCAACACAGCTTACTGCGGGCGGTGTGGCAATCCA 469
 QY 101 GlnThrIleuIleuThrAlaAlaHisCYsArgIleValValPheArgValArgLeuGlyHis 120
 DB 470 CAGTGGCTGTCAACGCGCCCACTGCAGAAAGAAATTTTCAGAGTCCGCTTCGCGCAC 529
 QY 121 TyrSerIleuSerProValTyrGlnSerGlyValGlnIleuMetPheGlnGlyValIleuSerIle 140
 DB 530 TACTCCCTGTCAACAGTTATGATCTGGGACAGATGTTCCAGGGGGTCAAAATCCATC 589
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspIleuIleuValIleuAsn 160
 DB 590 CCCACCTGTGCTACTCCACCTGGCACTTACGACCTCATGCTCATCAATCAATCAAC 649
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
 DB 650 AGAAGAAATTCGTCCCACTAAAGATGTCAGACCCATCAAGCTCTCTCATTTGCTCTCT 709
 QY 181 AlaGlyThrIleCYsIleuValIleuSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
 DB 710 GCTGGGCAAAAGTGTGTGTGCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCATTC 769
 QY 201 ProIleValIleuGlnCYsIleuAsnIleSerValIleuSerGlnIleuArgCYsGlyIuAspAla 220
 DB 770 CTTAAGGCTCTCAAGTCTTGAATATCAAGGTCTTAAGTCAAGAAAGTGCAGAGATGT 829
 QY 221 TyrProArgGlnIleAspAspThrMetPheCYsAlaGlyAspIleValArgAspSer 240
 DB 830 TACCCAGACAGATATGATGACACATGTTCTTGGCCGGTGCACAAACAGGTATAGATCC 889
 QY 241 CysGlnGlyAspSerGlyIleProValValCYsAsnGlySerIleuGlnIleuValIleuSer 260
 DB 890 TGCAGAGGTATCTGGGGGGGCTGTGTGTGCAATGGCTCCCTGCAGGGAGCTCGTGTCC 949
 QY 261 ThrGlyAspTyrProCYsAlaArgProAsnArgProGlyValIleuThrAsnIleuCYsIleu 280
 DB 950 TGGGAGATTACCTTGTGTGCTGGCCCAACAGACCGGTGTCTTACAGAACCTCTGCAAG 1009
 QY 281 PheThrIleTyrIleGlnIleuThrIleGlnAlaAsnSer 293
 DB 1010 TTACCAAGTGAATCCAGAAACCATCCAGGCCAATCC 1048
 RESULT 14
 ADA40367
 ID ADA40367 standard; cDNA; 1490 BP.
 XX
 AC ADA40367;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein encoding cDNA.
 XX
 KW Human; secreted protein; cancer; hyperproliferative disorder;
 KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
 KW anaemia; allergic reaction; asthma; cardiovascular disorder;
 KW wound healing; cytostatic; immunosuppressive; neotropic; neuroprotective;
 KW antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory;
 KW vulnery; cardiac; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2002102993-A2.
 XX
 XX 27-DEC-2002.

XX 19-MAR-2002; 2002WC-US008123.
 PF
 XX 21-MAR-2001; 2001US-0277340P.
 PR 19-UTL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI, 2003-175238/17.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating cancer or other hyperproliferative disorder,
 PT asthma, allergies or AIDS.
 PT
 XX
 PS Claim 9; SEQ ID NO 749; 3205bp; English.
 XX
 CC The invention relates to novel genes ADA39629-ADA40565 and proteins
 CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
 CC creating or ameliorating medical conditions e.g. by protein or gene
 CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
 CC fragments, and agonists or antagonists that bind to the polypeptide are
 CC useful for preparing a diagnostic or pharmaceutical composition for
 CC diagnosing or treating cancer or other hyperproliferative disorder. The
 CC polypeptides and nucleic acid molecules are also useful for detecting,
 CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
 CC or other hyperproliferative disorders including neoplasms, autoimmune
 CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
 CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
 CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
 CC thrombocytopenia), allergic reactions including asthma or eczema,
 CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
 CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
 CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
 CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
 CC fungal or viral infections including HIV/AIDS), or wound healing and
 CC disorders of epithelial cell proliferation. The nucleic acids are also
 CC useful for chromosome identification, radiation hybrid mapping or long-
 CC range restriction mapping, as molecular weight markers, or as
 CC hybridization or diagnostic probes. The polypeptides and antibodies are
 CC useful for providing immunological probes for differential identification
 CC of the tissues immunohistochemistry assays. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1490 BP; 323 A; 486 C; 373 G; 307 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1,46e-135 Length: 1490
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-936-271B-14 (1-293) x ADA40367 (1-1490)
 QY 1 MetAlaThrAlaArgProProlTpmActTPValLeuCyAlaLeuIleThrAlaLeuLeu 20
 DB 182 ATGGCTACACGACAGACCCCTCGATGTGGGTCTCTGCTCTGATCAGCTTGTCTT 241
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCyAspHisProSerAsn 40
 DB 242 CTGGGGGTACACAGAGCATGTTCTCGCAACATGATGTTCTCTGTGACCACTCTTAAC 301
 QY 41 ThrValProSerGlySerArgGlnAspLeuGlyValGlyValAlaGlyValAspAlaAsp 60
 DB 302 ACCGTGCTCTGTGGAGCAACGAGACTGGGAGCTGGGGCCGGGGAAGACGCCGCTCG 361

QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
 DB 362 GATTGACAGACGAGCGCATCATCATGATCGACTGGCATATGCACCCAGCCGTGG 421
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuIleTyrGlyValAlaValLeuValHisPro 100
 DB 422 CAGGCCGCGCTTGTCTAAGGCCCAACGAGCTCTACTGGGGCGGTGTGTGTGATCCA 481
 QY 101 GlnTPLeuLeuThrAlaAlaHisCyAspArgGlyValAlaPheArgValArgLeuGlyHis 120
 DB 482 CAGTGGCTCTCTACGCGCCGCCCATCTGCAGAGAGAGAGATTTTACAGATCCGTCTCGGCCAC 541
 QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnIleValIleSerIle 140
 DB 542 TACTCCCTGTACCAAGTTTATGATCTGGGACGAGATGTTCCAGGGGTCAATTCATTC 601
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
 DB 602 CCCACCCCTGGCTACTCCACCCCTGGCCACTTACGACCTCATCTCATCAACTGAAAC 661
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
 DB 662 AAGAGATTGTCCTCACTAAGATGTCAGACCCATCAACGTTCTCTCATTTGCTCTT 721
 QY 181 AlaGlyThrIleCyAlaLeuValSerGlyTyrGlyThrThrIleYserProGlnValHisPhe 200
 DB 722 GCTGGGACAAATGCTGTGGTGTCTGTGGGGGACACCAAGAGCCCAAGTGCACCTTC 781
 QY 201 ProIleValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
 DB 782 CCTAAGTCTCTCAGTGTCTGAATTCAGCGCTCAAGCAAGAAAGGCGCAGAGATGCT 841
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleAlaGlyValArgAspSer 240
 DB 842 TACCCGAGACATATGATGACACCATGTTCTGCCCGGTCAGAAAGCAGGTGAGACTCC 901
 QY 241 CysGlnGlyAspSerGlyValProValValCyAsnGlySerLeuGlnIleValSer 260
 DB 902 TTCAGGGGATTTCTGGGGGGCTGTGGTCTGATGCTCCCTCAGGAGACTCGTGTCC 961
 QY 261 TrpGlyAspTyrProCyAlaArgProAsnAspProGlyValTyrThrAsnLeuCyGlyAs 280
 DB 962 TGGGAGATTTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTTGCAG 1021
 QY 281 PheThrIleTyrIleGlnIleThrIleGlnIleAlaAsnSer 293
 DB 1022 TTCACCAAGTGGATCCAGAAACATTCAGGCCCACTCC 1060
 RESULT 15
 ADA56532
 ID ADA56532 standard; DNA; 1490 BP.
 AC XX
 AC ADA56532;
 DT 20-NOV-2003 (first entry)
 XX XX
 DE Gene encoding human secreted protein #257.
 XX XX
 KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
 KW cytostatic; cerebroprotective; neuroprotective; nocitropic;
 KW cardiovascular; antiarteriosclerotic; gene therapy;
 KW human secreted protein; immune disorder; inflammation;
 KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KW triple helix formation; antisense gene therapy; forensic biology; ds;
 KW gene.
 OS Homo sapiens.
 XX XX
 XX PN WO2002102994-A2.
 XX XX

PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002MO-US008278.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0311287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM,
 DR WPI; 2003-167512/16.
 XX P-PSDB; ADA57425.
 DR
 XX
 PT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or
 PT neurodegenerative disorders.
 XX
 PS Claim 21; SEQ ID NO 721; 1754bp; English.
 XX
 CC The invention relates to 592 new human secreted polypeptides useful for
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
 CC conditions, respiratory disorders, cancers, CNS disorders, or
 CC neurodegenerative disorders, or polypeptides comprising an amino acid
 CC sequence at least 95% identical to the new sequences. The polypeptides,
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic
 CC acids encoding the polypeptides, agonists or antagonists that binds to
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
 CC compositions for diagnosing, treating or preventing an e.g. immune
 CC disorder, inflammatory conditions (e.g. inflammatory bowel disease,
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 CC polynucleotides are useful for chromosome identification, chromosome
 CC mapping, for controlling gene expression through triple helix formation
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals
 CC from minute biological samples, in forensic biology, and as hybridization
 CC probes. The polypeptides are useful for as molecular weight markers on
 CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 CC gels, to raise antibodies, for testing biological activities, and for
 CC treating or preventing neural disorders, immune system disorders,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds
 CC to a gene encoding one of the polypeptide of the invention. Note: The
 CC sequence data for this patent did form part of the printed specification,
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 1490 BP; 323 A; 486 C; 373 G; 307 T; 0 U; 1 Other;
 XX
 Alignment Scores:
 Pred. No.: 1 46e-135 Length: 1490
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-936-271b-14 (1-293) x ADA56532 (1-1490)
 QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
 DB 182 ATGGCTACAGACAGACCCCTGATGTGGTGTCTGTCTGATCAGACGCTTCTT 241
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
 DB 242 CTGGGGGTCAAGAGCATGTTCGCAACATGATGTTCCTGTGACCAACCCCTTAC 301
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyIuAspAlaArgSer 60

DB 302 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAAGACGCCGCTCG 361
 QY 61 AspAspSerSerArgIleIleAlaGlySerAspCysAspMetHisThrGlnProTrp 80
 DB 362 GATGACAGACAGACCGCATCATCAATGATCCGACGTGCGATATGACACCCGCGTGG 421
 QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrosGlyValAlaValLeuValHisPro 100
 DB 422 CAGCGCGCTGTGTGCTAAGGCCCAACAGCTTACTGCGGGCGGTGTGTCATCCA 481
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
 DB 482 CAGTGGCTGTCAAGCGCCGCTCAGAGAAAGATTTCAGATGCTCTCGGCGAC 541
 QY 121 TyrSerLeuSerProValTyrosGlyGlnGlnMetPheGlnGlyValLysSerTle 140
 DB 542 TACTCCCTGTCAACAGATTATGAACTGGGACAGATGTTCCAGGGGGTCAATCCATC 601
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLysLeuAsn 160
 DB 602 CCCACCCCTGGCTACTCCACCTGGCCACTTACAGACTCATGCTCATCAACTGAAC 661
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
 DB 662 AGAAGAAATGTCTCCACTAAGATGTCAGACCATCAAGCTCTCTCTCATTTGCTCT 721
 QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
 DB 722 GTGGGACAAAGTGGTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCAAGTGCACTTC 781
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
 DB 782 CTPAAGTCTCTCAAGTCTTGAATATCAAGGTCTAAGTCAAGAAAGTCCGAGATGCT 841
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
 DB 842 TACCCAGACAGATATGATGACCATGTTCTGCGCCGTCACAAAGAGATAGACTCC 901
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
 DB 902 TGCACAGGTATCTGGGGGGCTGTGTGTGCAATGGCTCCCTGCAGGAGATCTGTCTCC 961
 QY 261 TrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrrThrAsnLeuCysLys 280
 DB 962 TGGAGAAATTAACCTTGTGCTCGCCCAACAGACCGGGGTGTTCACAGAACTCTGCAAG 1021
 QY 281 PheThrLysTrpIleGlnGlnIuThrIleGlnAlaAsnSer 293
 DB 1022 TTCACCAAGTGAATCCAGAAACATCCAGGCCAATCC 1060

Search completed: February 26, 2005, 12:58:47
 Job time : 617 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: February 26, 2005, 12:32:30 / Search time 204 Seconds

(without alignments)
2350.143 Million cell updates/sec

Title: US-09-936-271B-14
Perfect score: 1608
Sequence: 1 MATARPPMVMVLCALITALL.....VYTNLCKFTKMIQTIQANS 293

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV=xlh
-Q=/cg2_1/USPTO.spool/US0936271/runat_23022005_130127_27287/app_query.fasta_1.455
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum2 -TRANS=human40.ccd
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NOBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US0936271.QCGN_1.1_66@runat_23022005_130127_27287 -NCPV=6 -ICPV=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued_Patents_NA.*
1: /cg2_5/ptodata/1/ina/5A.COMB.seq.*
2: /cg2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cg2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cg2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cg2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cg2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|-----------------------|--------------------|
| 1 | 1608 | 100.0 | 1499 | 4 US-09-509-908-1 | Sequence 1, Appl1 |
| 2 | 1595 | 99.2 | 1476 | 2 US-08-824-874-2 | Sequence 2, Appl1 |
| 3 | 1595 | 99.2 | 1476 | 3 US-09-210-084-2 | Sequence 2, Appl1 |
| 4 | 1595 | 99.2 | 1476 | 4 US-09-764-762-2 | Sequence 2, Appl1 |
| 5 | 1479 | 92.0 | 1143 | 4 US-09-949-016-1841 | Sequence 1841, Ap |
| 6 | 1403 | 87.3 | 1504 | 3 US-09-280-116-1 | Sequence 1, Appl1 |
| 7 | 1002 | 62.3 | 10818 | 4 US-09-949-016-13583 | Sequence 13583, A |
| 8 | 725 | 45.1 | 735 | 3 US-09-602-877A-94 | Sequence 94, Appl1 |
| 9 | 671 | 41.7 | 765 | 3 US-09-439-313-524 | Sequence 524, App |
| 10 | 671 | 41.7 | 765 | 4 US-09-636-215-524 | Sequence 524, App |
| 11 | 671 | 41.7 | 765 | 4 US-09-685-166A-524 | Sequence 524, App |
| 12 | 671 | 41.7 | 765 | 4 US-09-679-426-524 | Sequence 524, App |

| | | | | | |
|----|-------|------|------|----------------------|--------------------|
| 13 | 671 | 41.7 | 765 | 4 US-09-759-143-524 | Sequence 524, App |
| 14 | 671 | 41.7 | 765 | 4 US-09-651-236-524 | Sequence 524, App |
| 15 | 654.5 | 40.7 | 1052 | 3 US-09-386-642-110 | Sequence 10, Appl |
| 16 | 652.5 | 40.6 | 1146 | 4 US-09-205-258-247 | Sequence 247, App |
| 17 | 652.5 | 40.6 | 1166 | 3 US-08-944-483-7 | Sequence 7, Appl1 |
| 18 | 650.5 | 40.5 | 833 | 2 US-08-790-137-2 | Sequence 8, Appl1 |
| 19 | 650.5 | 40.5 | 1192 | 3 US-08-944-483-8 | Sequence 189, App |
| 20 | 650.5 | 40.5 | 1292 | 4 US-09-205-258-189 | Sequence 189, App |
| 21 | 650.5 | 40.5 | 1314 | 3 US-09-025-059-2 | Sequence 2, Appl1 |
| 22 | 641.5 | 39.9 | 1343 | 4 US-09-618-259-72 | Sequence 72, Appl |
| 23 | 641.5 | 39.9 | 1360 | 4 US-09-618-259-6 | Sequence 6, Appl1 |
| 24 | 640 | 39.8 | 1350 | 4 US-09-636-215-616 | Sequence 616, App |
| 25 | 640 | 39.8 | 1350 | 4 US-09-685-166A-616 | Sequence 616, App |
| 26 | 640 | 39.8 | 1350 | 4 US-09-679-426-616 | Sequence 616, App |
| 27 | 640 | 39.8 | 1350 | 4 US-09-759-143-616 | Sequence 616, App |
| 28 | 639.5 | 39.8 | 1350 | 4 US-09-651-236-616 | Sequence 616, App |
| 29 | 639.5 | 39.8 | 944 | 3 US-09-008-271A-19 | Sequence 19, Appl1 |
| 30 | 638.5 | 39.7 | 994 | 3 US-09-070-526-1 | Sequence 1, Appl1 |
| 31 | 638.5 | 39.7 | 1253 | 4 US-09-949-016-2295 | Sequence 2295, Ap |
| 32 | 638 | 39.7 | 871 | 1 US-08-744-026-2 | Sequence 2, Appl1 |
| 33 | 638 | 39.7 | 871 | 2 US-09-102-732-2 | Sequence 2, Appl1 |
| 34 | 638 | 39.7 | 871 | 3 US-09-261-767-2 | Sequence 2, Appl1 |
| 35 | 638 | 39.7 | 871 | 3 US-08-969-987-7 | Sequence 7, Appl1 |
| 36 | 638 | 39.7 | 1248 | 3 US-09-020-956-171 | Sequence 171, App |
| 37 | 638 | 39.7 | 1248 | 3 US-09-030-607-171 | Sequence 171, App |
| 38 | 638 | 39.7 | 1248 | 3 US-09-439-313-171 | Sequence 171, App |
| 39 | 638 | 39.7 | 1248 | 3 US-09-352-616A-171 | Sequence 171, App |
| 40 | 638 | 39.7 | 1248 | 3 US-09-232-149A-171 | Sequence 171, App |
| 41 | 638 | 39.7 | 1248 | 4 US-09-159-812-171 | Sequence 171, App |
| 42 | 638 | 39.7 | 1248 | 4 US-09-636-215-171 | Sequence 171, App |
| 43 | 638 | 39.7 | 1248 | 4 US-09-685-166A-171 | Sequence 171, App |
| 44 | 638 | 39.7 | 1248 | 4 US-09-115-453-171 | Sequence 171, App |
| 45 | 638 | 39.7 | 1248 | 4 US-09-688-489-171 | Sequence 171, App |

ALIGNMENTS

RESULT 1

US-09-509-908-1
Sequence 1, Application US/09509908
Patent No. 6589770

GENERAL INFORMATION:

APPLICANT: The Procter & Gamble Company, N/A N/A
TITLE OF INVENTION: A Protease
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: T. David Reed
STREET: 5299 Spring Grove Avenue
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZIP: 45217-1087

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/509,908
FILING DATE: 28-Feb-2002
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Reed, T David
REGISTRATION NUMBER: 32,931
REFERENCE/DOCKET NUMBER: AA-264P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 513-627-7025
TELEFAX: 513-627-6333

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```
/
/
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 291..1172
/
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 489..1172
/
/ NAME/KEY: 5'UTR
/ LOCATION: 1..230
/
/ NAME/KEY: 3'UTR
/ LOCATION: 1173..1499
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-509-908-1

Alignment Scores:
Pred. No.: 6.98e-164 Length: 1499
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-936-271B-14 (1-293) x US-09-509-908-1 (1-1499)

QY 1 MetAlaThrAlaArgProPrtPmetTrrValLeuCyAlaLeuIleThraAlaLeu 20
DB 291 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGTCTGATCAAGCCCTTGT 350
QY 21 LeuGlYValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 351 CTGGGGGTCAAGAGCATGTTCTGCCCAAGATGTTTCCGTGACCAACCCCTTAC 410
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyAlaGlyValAspAlaArgSer 60
DB 411 ACCGTGCTCTGGGGCAACCAAGCACTGGGAGCTGGGGCCGGGGAAGACGCCGGTGC 470
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrr 80
DB 471 GATGACAGCAGCAGCGCATCATCATGATCCGACTGCGATATGCAACCCAGCCGTGG 530
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrrCyAsGlyAlaValLeuValHisPro 100
DB 531 CAGGCGCGCTGTGTGCTAAGGCCCAACAGCTTACTCGGGGCGGTGTGGTGCATCCA 590
QY 101 GlnTrrPleuLeuThrAlaAlaHisCyAsGlyValValPheArgValArgLeuGlyHis 120
DB 591 CAGTGGCTGCTCAAGCGCCGCCCTCAGAGGAGAAAGTTTTCAGAGTCCGTCTCGGCCAC 650
QY 121 TyrSerLeuSerProValTrrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 651 TACTCCCTGTACACAGTTTATGATCTGGGACAGATGTTCCAGGGGGTCAATATCCATC 710
QY 141 ProHisProGlyTrrYserHisProGlyHisSerAsnAspLeuMetLeuIleValAsn 160
DB 711 CCCACCCCTGGCTACTCCCACTAAGATGTCAGACCCATCAACGTCCTCTCTGCTCCCTC 770
QY 161 ArgAlaGlnIleArgProThrIlyAspValArgProIleAsnValSerSerHisCyAsProSer 180
DB 771 AGAAGATTCCTCCCACTAAGATGTCAGACCCATCAACGTCCTCTCTGCTCCCTC 830
QY 181 AlaGlyThrIlyAspValSerGlyTrrGlyThrThrIlySerProGlnValHisPhe 200
DB 831 GCTGGGACAAAGCTTGTGTGTGGCTGGGGGACCAACCAAGGCCCCCAAGTGCATTC 890
QY 201 ProIlyValLeuGlnCyAsuAsnIleSerValLeuSerGlnIlyAsArgCyAsGluAspAla 220
DB 891 CCAAGGTCTCTCAAGTGTGATATCAAGGTCTTAAGTCAAGAAAGTCCAGAGATGCT 950
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIlyAsAlaGlyArgAspSer 240
DB 950
```

```
DB 951 TACCCGAGACAGATAGATGACACCATGTTCTGCCCGGTGACAAAGCAGTAGAGACTCC 1010
QY 241 CyGlnGlnIlyAspSerGlyValIlyProValValCyAsnGlySerLeuGlnIlyLeuValSer 260
DB 1011 TGCAGGGTATTCGAGGGGCGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1070
QY 261 TrrGlyAspTrrProCyAsAlaArgProAsnArgProGlyValTrrThrAsnLeuCyAsIly 280
DB 1071 TGGGGAGATTACCTTGTGCCCGGCCCAACAGACCGGGGTCTTACACGAACCTTGCAG 1130
QY 281 PheThrIlySTrrIleGlnIlyThrIleGlnAlaAsnSer 293
DB 1131 TTCACCAAGTGTATCCAGGAACCATCCAGGCCAACTCC 1169

RESULT 2
US-08-824-874-2
/ Sequence 2, Application US/08824874
/ Patent No. 5962300
/ GENERAL INFORMATION:
/ APPLICANT: Hillman, Jennifer L.
/ APPLICANT: Lal, Preeti
/ TITLE OF INVENTION: NOVEL KALLIKREIN
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/824,874
/ FILING DATE: Filed Herewith
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1476 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KERANOT02
/ CLONE: 820694
US-08-824-874-2

Alignment Scores:
Pred. No.: 1.73e-162 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: Gaps: 0

US-09-936-271B-14 (1-293) x US-08-824-874-2 (1-1476)

QY 1 MetAlaThrAlaArgProPrtPmetTrrValLeuCyAlaLeuIleThraAlaLeu 20
DB 277 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGTCTGATCAAGCCCTTGTCTT 336
```

```

QY      21  LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerAsn 40
DB      337  CTGGGGGTCACAGAGCATGTTCTGCCAACAATATGTTCTCTGTGACCAACCCCTCTAAC 396
QY      41  ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyValAspAlaArgSer 60
DB      397  ACCGTGCTCTGTGGAGCAACCAAGACCTGGGAGCTGGGGCGGGGAGAACGCCGGTTCG 456
QY      61  AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB      457  GATGACAGCAGCAGCCCATCATCATGATGATCCAGCTGCGGATATGACACCCACCGCTGG 516
QY      81  -GlnAlaAlaLeuLeuLeuLeuArgProAsnGlnLeuTyrcysGlyValAlaValLeuValHisPr 100
DB      517  GCAGGGCGCGCTGTGCTAGAGGCCCAACCAAGCTTACTGCGGGCGGGTGTGGTGCATCC 576
QY      100  OGIntPrLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
DB      577  ACAGTGCTGCTCACCGGCCGCCCATCTGACGAGAAAGATTTCAGAGTCCGCTCGGCCA 636
QY      120  sTyrsSerLeuSerProValTyrglySerGlyGlnGlnMetPheGlnGlyValLysSer11 140
DB      637  CTACTCCCTGTCAACAGTTATGATCTGGGACAGATGTTCCAGGGGGTCAAAATCAT 696
QY      140  eProHisProGlyTyrsSerHisProGlyHisSerAspLeuLeuLeuLeuLeuLeuLeu 160
DB      697  CCCCACCTGGCTACTCCACCTGGCCCATCTTAACGACTCATGCTCATCAACCTGA 756
QY      160  nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180
DB      757  CAGAGAAATTCCTCCCACTAAAGATGTCAGACCATCACTCTCTCTCATCTCTCTC 816
QY      180  rAlaGlyThrLysCysLeuValSerGlyTyrglyThrThrLysSerProGlnValHisPh 200
DB      817  TGCTGGGACAAAGCTGTGGTCCGGCTGGGGGACAAACCAAGACCCCAAGTGCACTT 876
QY      200  eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlnuAspAl 220
DB      877  CCTTAAGCTCTCCAGTGTGTAATATCAGCGTCTAAGTCAAGAAAGTCCGAGATGC 936
QY      220  sTyrsProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysValGlyArgAspSe 240
DB      937  TTACCCAGACAGATAGTATGACACCATGTTCTGGCCGGTGCACAAAGCAGGTAGAGCTC 996
QY      240  rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
DB      997  CTGCCAAGGTGATTTCTGGGGGGCTGTGTCTGCAATGGCTCCTGCAGGGGACTCGTGT 1056
QY      260  rTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrcysThrAsnLeuCysLy 280
DB      1057  CTGGGGAGATTATCTGTGTGCTGGCCCAACAGACCGGGGTCTTAACGAACCTCTGCAA 1116
QY      280  sPheThrLysTrpIleGlnGlnuThrIleGlnAlaAsnSer 293
DB      1117  GTTCAACCAAGTGAATCCAGGAACCATCCAGGCCAATCTCC 1156

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-09-210-084-2
Alignment Scores:
Pred. No.: 1 73e-162 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 3 Gaps: 0
US-09-936-271b-14 (1-293) x US-09-210-084-2 (1-1476)
QY      1  MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
DB      277  ATGGCTAAGCAGCAAGACCCCTGGATGGGTGCTCTGTCCTGTATCACAGCCTTGCTT 336
QY      21  LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerAsn 40
DB      337  CTGGGGGTCAACAGCATGTTCTGCCAACAATGATGTTCTGTGACCAACCCCTCTAAC 396
QY      41  ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyValAspAlaArgSer 60
DB      397  ACCGTGCTCTGTGGAGCAACCAAGACCTGGGAGCTGGGGCGGGGAGAACGCCGGTTCG 456
QY      61  AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB      457  GATGACAGCAGCAGCCCATCATCATGATGATCCAGCTGCGGATATGACACCCACCGCTGG 516
QY      81  -GlnAlaAlaLeuLeuLeuLeuArgProAsnGlnLeuTyrcysGlyValAlaValLeuValHisPr 100
DB      517  GCAGGGCGCGCTGTGCTAGAGGCCCAACCAAGCTTACTGCGGGCGGGTGTGGTGCATCC 576
QY      100  OGIntPrLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
DB      577  ACAGTGCTGCTCACCGGCCGCCCATCTGACGAGAAAGATTTCAGAGTCCGCTCGGCCA 636
QY      120  sTyrsSerLeuSerProValTyrglySerGlyGlnGlnMetPheGlnGlyValLysSer11 140
DB      637  CTACTCCCTGTCAACAGTTATGATCTGGGACAGATGTTCCAGGGGGTCAAAATCAT 696
QY      140  eProHisProGlyTyrsSerHisProGlyHisSerAspLeuLeuLeuLeuLeuLeuLeu 160
DB      697  CCCCACCTGGCTACTCCACCTGGCCCATCTTAACGACTCATGCTCATCAACCTGA 756
QY      160  nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180

```

Db 757 CAGAGAAATTCGTCCTCAAGATGTCAGACCCATCAAGTCTCCCTCATTTGCTCC 816
QY 180 rlaaglythlyscyleuValSerGlyTTPGlyThrThlySerProGlnValHisph 200
Db 817 TCGTGGACAAAGTGTCTGTGTCCGCTGGGGGACAAACCAAGCCCTCAAGTGCATT 876
QY 200 ePcylValleuGlnCysleuAniIleSerValleuSerGlnLysArgCysGlnAsp 220
Db 877 CCTTAAGTCTCCAGTGTCTTGAATATCAGCTGTCTAGTCAGAAAGGTGCCAGATGC 936
QY 220 alyrProArgGlnIleAspAspThrePheCysAlaGlyAspLysAlaGlyArgAsp 240
Db 937 TTACCCGACAGACATGATGACACACATGTTCTGCGCCGTCGACAAAGCAGTAGACCTC 996
QY 240 rCysGlnGlyAspSerGlyLysProValValCysAsnGlySerLeuGlnLysValSe 260
Db 997 CTGCCAGGGGATCTGGGGGGCTGTGTCTCAATGCTCTCCCTGCAGGAGCTCGTGC 1056
QY 260 rTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysly 280
Db 1057 CTGGGAGATTACCTTTGTCTGCCGCCCAACACACCGGGTGTCTACACGAACTTTCGAA 1116
QY 280 spherhlyeTrrIleGlnGlnThrIleGlnAlaAsnSer 293
Db 1117 GTTCAACCAAGTGCATCCAGGAACCATCCAGGCCAATCC 1156

RESULT 4

US-09-764-762-2
Sequence 2, Application US/09764762
Patent No. 6472195
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/764,762
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,084
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 820694
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Alignment Scores:

| Pred. No.: | 1,736-162 | Length: | 1476 |
|------------------------|-----------|---------------|------|
| Score: | 1595.00 | Matches: | 293 |
| Percent Similarity: | 99.66% | Conservative: | 0 |
| Best Local Similarity: | 99.66% | Mismatches: | 0 |
| Query Match: | 99.19% | Indels: | 1 |
| DB: | 4 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x US-09-764-762-2 (1-1476)

QY 1 MetAlaThrAlaArgProTrrPmetTrrValleuCysAlaLeuIleThrAlaLeu 20
Db 277 ATGGCTACGACGAACCCCTCGATGTGGGTGTCTGTGTATCAACCTTGTCTT 336
QY 21 LeuGlyValThrGlnHisValleuAlaAsnAspValSerCysAspHisProSerAn 40
Db 337 CTGGGGGTCAAGAGATGTCTCGCCAAATGATGTTCTGTGACCACTCTTAC 396
QY 41 ThrValProSerGlySerAnGlnAspLeuGlyValAlaGlyLysAspArgSer 60
Db 397 ACCGTGCTCTGGAGACACCAAGACTGGAGCTGGGGCCGGGAGACCCCGGTG 456
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrr 80
Db 457 GATGACAGCAGCAGCCGATCATGATGATCCGATGATATGACACCCAGCCGTGG 516
QY 81 -GlnAlaAlaLeuLeuArgProAnGlnLeuTyrCysGlyAlaValleuValHisPr 100
Db 517 GCAGGCCCGCTGTGTGTCTAAGGCCCAACCAAGCTTACTGGGGGGGTGTGTGATCC 576
QY 100 OGlnTrrPleuLeuTrrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 577 ACGTGGCTGTCTACGCGCCGCTGACGAGAAAGTTTCAAGTCCGTCTGCGCA 636
QY 120 sTyrSerLeuSerProValTyrGlnSerGlyGlnMetPheGlnGlyValLysSer11 140
Db 637 CTACCTCCCTGACACGATTATGAATCTGGGACAGCAAGTTTCCAGGGGGTCAAACTCAT 696
QY 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
Db 697 CCCCCACCTGTGTACTCCACCTGGCCACTCTTAACGACTCAAGCTCATCAACTGAA 756
QY 160 nArgArgIleAspProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180
Db 757 CAGAGAAATTCCTCCATTAAGATGTCAAGCCATCAAGCTCTCTCATTTGCCCTC 816
QY 180 rlaaglythlyscyleuValSerGlyTTPGlyThrThlySerProGlnValHisph 200
Db 817 TCGTGGACAAAGTGTCTGTGTCCGCTGGGGGACAAACCAAGGCCCTCAAGTGCATT 876
QY 200 ePcylValleuGlnCysleuAniIleSerValleuSerGlnLysArgCysGlnAsp 220
Db 877 CCTTAAGTCTCCAGTGTCTTGAATATCAGCTGTCTAATTCAGAAAGGTGGAGATGC 936
QY 220 alyrProArgGlnIleAspAspThrePheCysAlaGlyAspLysAlaGlyArgAsp 240
Db 937 TTACCCGACAGACATGATGACACATGTTCTGCGCCGTCGACAAAGCAGTAGACCTC 996
QY 240 rCysGlnGlyAspSerGlyLysProValValCysAsnGlySerLeuGlnLysValSe 260
Db 997 CTGCCAGGGGATCTGGGGGGCTGTGTCTGCAATGCTCTCCCTGCAGGAGCTCGTGC 1056
QY 260 rTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysly 280
Db 1057 CTGGGAGATTACCTTTGTCTGCCGCCCAACACACCGGGTGTCTACACGAACTTTCGAA 1116
QY 280 spherhlyeTrrIleGlnGlnThrIleGlnAlaAsnSer 293
Db 1117 GTTCAACCAAGTGCATCCAGGAACCATCCAGGCCAATCC 1156

RESULT 5

US-09-949-016-1841


```
; Sequence 1841, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIORITY APPLICATION NUMBER: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; PRIORITY APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1841
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1841
Alignment Scores:
Pred. No.: 4,1e-150 Length: 1143
Score: 1479.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.98% Indels: 0
DB: Gaps: 0
US-09-936-271b-14 (1-293) x US-09-949-016-1841 (1-1143)
QY 25 GlnH1eValLeuAlaAnaaspValSerCyAspHisProSerAenThrValProSer 44
DB 6 GAGATGTTCTCGCCAAATGATGTTCTCGTGACACCCCTTAACAGCTGCTCT 65
QY 45 GlySerEnGlnIleuLeuGlyValAGlyValAGlyGluAspAlaAspSerAspSerSer 64
DB 66 GGGGCAACACAGACCTCGGAGCTGGGCGGGGAGAACCCCGCTGGATGACAGCAGC 125
QY 65 SerArgIleIleuAnGlySerAspCyAspMetHisThrGlnProTropGlnAlaAlaLeu 84
DB 126 AGCGGCAATCAATGATGATCGACTCGGATGACACCCGCTGGACGCGCTG 185
QY 85 LeuLeuArgProAenGlnLeuTyCyGlyValAlaValLeuValHisProGlnIleu 104
DB 186 TTGCTAAGGCCCAACAGCTCTACTGGGGGCGGTGTTGGATCCACAGTGGCTGCTC 245
QY 105 ThrAlaAlaHisCyAspArgValPheArgValArgLeuGlyHisTyrSerLeuSer 124
DB 246 ACCGCCCCCACTGACAGAAAGATTTCAGAGTCCGCTCGGCACTCTCCCTGCA 305
QY 125 ProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIleuValSerIleProHisProGly 144
DB 306 CCGATTATGATGTGGACAGACAGATGTCAGGGGGGCAATCAATCAATCCACCTGGC 365
QY 145 TyrSerHisProGlyHisSerAspAspLeuMetLeuIleuValLeuAspArgIleArg 164
DB 366 TACTCCCACTGGCCCACTCTAAGACCTCATGCTCATCAACAGAAAGAAATTCGT 425
QY 165 ProThrIleAspValAlaArgProIleAspValSerSerHisCyAspProSerAlaGlyThrLys 184
DB 426 CCCACTAAGATGTCAGACCCATCAAGCTCTCTCTCATGTCTCCCTGCTGGGACAAAG 485
QY 185 CyLeuValSerGlyTropGlyThrThrIleSerProGlnValHisPheProLysValLeu 204
DB 486 TGGTGGATGTGGCTGGGGACACCAAGACCCCAAGTGCATTCCTTAAGGCTCTC 545
QY 205 GlnCyLeuValHisSerValLeuSerGlnLysArgCyGluAspAlaTyrProArgGln 224
DB 546 CAGGCTTGATATCAAGGCTGCTAAGTCAAGAAAGGTGCGAGATGCTTACCGACAG 605
```

```
QY 225 IleAspAspThrMetPheCyAspAlaGlyAspIleValArgAspSerCyGlnGlyAsp 244
DB 606 ATGATGACACCATGTTCTGCGCGGTGACAAAGCATGATGACCTCCGCGGCTGAT 665
QY 245 SerGlyValProValValCyAspAnGlySerLeuGlnGlyLeuValSerTropGlyAspTyr 264
DB 666 TCGGGGGGGCTGTGCTGCTGCAATGAGCTCCCTGACAGGAGACTGCTGCGGGAGATTAC 725
QY 265 ProCyAlaArgProAenArgProGlyValTyrThrAsnLeuCyAlaSerPheThrIleTrop 284
DB 726 CCTTGTGCGCGCGCCCAACAGACCGGGTGTCTTACACGAACCTTGCAAGTTCAACAAGTGG 785
QY 285 IleGlnGluThrIleGlnAlaAspSer 293
DB 786 ATCCAGAAACATTCAGGCCAATCTCC 812
RESULT 6
US-09-280-116-1
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ. ID NOS: 268
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-1
Alignment Scores:
Pred. No.: 1,03e-141 Length: 1504
Score: 1403.00 Matches: 280
Percent Similarity: 93.77% Conservative: 6
Best Local Similarity: 91.80% Mismatches: 7
Query Match: 87.25% Indels: 12
DB: Gaps: 2
US-09-936-271b-14 (1-293) x US-09-280-116-1 (1-1504)
QY 1 MetAlaThrAlaArgProTropMetTropValLeuCyAlaLeuIleThrAlaLeuLeu 20
DB 277 ATGGCTACAGACAGACCCCTCGATGTGGGTCTGTGCTGTATCAACGCTTGCTT 336
QY 21 LeuGlyValThrGlnHisValIleuAla-AlaAnaaspValSerCyAspHisProSerArg 40
DB 337 CTGGGGGTCAACAGATGTTTCGCCAAACATGATGTTTCCTGTGACCAACCCCTCTAA 396
QY 40 nThrValProSerGlySer-AenGlnAspLeuGlyAla---GlyValAGlyGluAspAla 58
DB 397 CACCGTGGCTCTGTGGAGCAACACAGACTTGGGAACTGGGGCCCGGGGAAGAGGCC 456
QY 59 ArgSerAspAspSerSerSer-ArgIleIleAnGlySerAsp-CyAspMetHisThrG 78
DB 457 CGGTGGATACACAGACAGCCCGCATCATCATGATGATGCACTGGATATGACACCC 516
QY 78 IndProTropGlnAla---AlaLeuLeuValArgProAnGlnLeu--TyrCyGlyValAla 96
DB 517 AGCGGTGGCAGGCCCGCGCTGTGTTAAAGCCCAACCAAGTTCTTATTCGCGGGCGGT 576
QY 96 IleuValHisProGlnIleuLeuThrAlaAla-HisCyAspArgValPheArgVal 116
DB 577 GTTGGTGCATTCACAGAGGTGTTCAAGGGCCGCCCATGCAAGAAAGAAATTTTCAGAG 636
QY 116 AlaArgLeuGlyHisTyrSerLeuSerProValTyrGluSerGlyGlnGln-MetPheGln 135
DB 637 TCGGTCTCGGCCCATATTCCTCTGTCAACAGTTTATGATCTGGGACAGCAAGATGTTCCAG 696
```

| | | | |
|----|------|--|------|
| QY | 136 | GIYVallVesSerIleProHisProGlyThrSerHisProGlyHisSerAspAspLeuMet | 155 |
| Db | 697 | GGGTCAATCATCTCCACCCTGGGCTACTCCACCCTGGCCACTTAACGACTTATG | 756 |
| QY | 156 | LeuIleIysLeuAsnArgArgIleArgProThrIysAspValArgProIleAsnValSer | 175 |
| Db | 757 | CTCATCAAACTGAACAGAAAGAAATTCGTCCCACTAAAGATGTCAGACCATCAACTCTCC | 816 |
| QY | 176 | SerHisCysProSerAlaGlyIleThrIysCysLeuValSerGlyIleProGlyIleThrIysSer | 195 |
| Db | 817 | TCTCATGTCCTCTGCTGGAGCAAAAGCTGTGGTGTCTGGCTGGGGAGCAACCAAGC | 876 |
| QY | 196 | ProGlnValHisPheProIysValIleGlnCysLeuAsnIleSerValIleuSerGlnIys | 215 |
| Db | 877 | CCCCAAGGCACTTCCCTTAAGTCTCTCCAGTCTGTGAATATCACCGTGTCTTAAGTCAGAA | 936 |
| QY | 216 | ArgCysGluAspAlaIleIleProArgGlnIleAspAspThrMetPheCysAlaGlyAspIys | 235 |
| Db | 937 | AGGTCCGAGGATGCTTACCACGACAGATAGATGACACCATGTTCTCGCCGGGAGCAAA | 996 |
| QY | 236 | AlaGlyArgAspSerCysGlnGlyAspSerIleGlyProValIleCysAsnGlySerLeu | 255 |
| Db | 997 | GCAAGTACGACTCTCTGCCAGGAGGATTTCTGGGGGGCTGTGGTGTGTAATGGCTCTCCGTG | 1056 |
| QY | 256 | GlnGlyIleuValSerTyr-GlyAspIleIleProCysAlaArgProAsnArgProGlyValIle | 275 |
| Db | 1057 | CAGGAGACTCGTGTCTGGGGGAGATTAACCTTGTGCCCGGCCCAACAGACCGGGGTGTCTA | 1116 |
| QY | 275 | rThrAsnLeuIleCysIysPheThrIleIleTyrIleGlnGluThrIleGlnIleAsnSer | 293 |
| Db | 1117 | CACGAACTCTGCAGATTCACCAAGTGGATTCAGAAACCATTCAGGCGCAACTCC | 1171 |

RESULT 7

```

US-09-949-016-13583
; Sequence 13583, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13583
; LENGTH: 10818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13583

```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-------|
| Pred. No.: | 5.01e-97 | Length: | 10818 |
| Score: | 1002.00 | Matches: | 218 |
| Percent Similarity: | 44.13% | Conservative: | 0 |
| Best Local Similarity: | 44.13% | Mismatches: | 0 |
| Query Match: | 62.31% | Indels: | 276 |
| DB: | 4 | Gaps: | 2 |

US-09-936-271B-14 (1-293) X US-09-949-016-13583 (1-10818)

QY 25 GIHHISVALLLEUADAENHASPVALSERCYASPHIAPROSERATHHVALPROSER 44
DB 2006 GAGCATGTTTCGCCMACATGATGTTTCTGTGTACCAACCCCTTAAACCCGTGCTCT 20655
QY 45 Glyceranglinapleucliyalaciyalaciygulimupalatrgseraspaspserser 64

| | | | |
|----|------|--|------|
| Db | 2066 | GGAGCAACCAAGCACTGGAGCTGGGGCCGGGGAAGACCCCGGTCCGATACAGACAC | 2122 |
| QY | 65 | SerArgIleIleangIySeraspCyaaApwethIstnngInProtrpGlnAlaIleu | 84 |
| Db | 2126 | AGCGCATCATCAATGATCCGACTGGCATATGACACCCAGCCGTGGCAGGCCGCGCTG | 2185 |
| QY | 85 | LeuIleuAaGProbaenGlnLeuTyrcyaelIyAlaValIleuValHisProGlnItrPleuIleu | 104 |
| Db | 2166 | TTGCTAAAGCCCCMACCAAGCTCTACTGCGGGCGGTGGTGTGCATCCACAGTGGCTGCTC | 2244 |
| QY | 105 | ThraIaIaHisCyArGLyIeIsb----- | 112 |
| Db | 2246 | ACGCGCCGCTCACTGCAGGAAGAAGTGAAGTGGAGTTCCAAAGAGAGGTTGGTGGGAGC | 2304 |
| QY | 112 | ----- | 112 |
| Db | 2306 | GGGAAGTGGGGGTGGGGGTGGGGAAAGTGGGGGTGGGGGTGTCAATGAGGTGAGGGGTGT | 2366 |
| QY | 112 | ----- | 112 |
| Db | 2366 | GGGACGGGGAAAGTGGGGTGGGGGTGTCAATGAGGTGAGGGTGTGGGGATGGGTTG | 2422 |
| QY | 112 | ----- | 112 |
| Db | 2426 | GGGATGTGGAGCAGAGAGAGGTGAGTTGGGGATAGACTTAAGATGAGATTGTTCGGG | 2486 |
| QY | 112 | ----- | 112 |
| Db | 2486 | GGAGCAAGGTGGAGAGATGAGGTGTGGAAGGGGAGAGTGTGTGATGGAAATGGGAAG | 2544 |
| QY | 112 | ----- | 112 |
| Db | 2546 | AGCCAAGATGGGTGTGATTTGGGGTTAGAGCATATATTTGTGAAATGCTTTGGGATGG | 2606 |
| QY | 112 | ----- | 112 |
| Db | 2606 | AGTGGAAATTGGGATTGGCTTTAGAAATTGGGGGTGGGAAATCGGGCTGGGGTGGAAA | 2666 |
| QY | 112 | ----- | 112 |
| Db | 2666 | TGAAGATGCAATGAGATAGGGTGTGAGTTTGGGAGACAGATATAGAAATAGAGATGGGAT | 2722 |
| QY | 112 | ----- | 112 |
| Db | 2726 | TGAGATTTTGGGTGGGGTGTGAGATGTTGGAATTTGGGCTTGAGAAATCATATGGTGAATG | 2786 |
| QY | 112 | ----- | 112 |
| Db | 2786 | GCTTCTGGGTAGGGAAAGAAATTAGGGTGTGGAAATGGGATGGGTTGGAATGTGACTGGG | 2844 |
| QY | 112 | ----- | 112 |
| Db | 2846 | ATGGGGACAGAGCATGGGATTTGGAGACCAAGAGGAGATTGAGGATGGTTTGGGACCGGGG | 2904 |
| QY | 112 | ----- | 112 |
| Db | 2906 | GTGGGATGGGGGTGGGGCTGGGGCTGGGTGTGGGATTTGGGATGGCGTGGACGTGGAG | 2966 |
| QY | 113 | -----ValPheArgValArgLe----- | 118 |
| Db | 2966 | ATAGAGATCAAGGTTGTGTGACCTGCCCCATCTTCTCAGAGATTTTCAAGTCCGCTT | 3022 |
| QY | 118 | uGIyHisItySerIeuserProvalItyrGIuserGIyGlnIImetPheGlnIyValIy | 138 |
| Db | 3026 | CGGCGCACTACTCCCTGTACCAAGTTTATGAATCTGGGACAGAGATGTTTCCAGGGGGTCA | 3086 |
| QY | 138 | sSerIleProHisProGIyTySerHisProGIyHisSerIeAsnApleuMetLeuIleIy | 158 |
| Db | 3086 | ATTCATCCCCCAACCTGGCTACTCCCAACCTGGGCACTCTAAGCACTCATGTCTCATCA | 3144 |
| QY | 158 | sIeuaAsnArgIleArgIleArgProthIyAspValArgProIleAsnValSerSerHisCy | 178 |

Db 3146 ACTGACAGAGAAATTCCTCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATTTG 3205
Qy 178 sproseralaglythrlyscysleuvalserglytrpglythrthrlyserprogin-- 197
Db 3206 TCCCTGTGCTGGGACAAAGTCTGTGTCTGGGGGACAAAGAGAGAGAGAGAGAGAGAG 3265
Qy 197 ----- 197
Db 3266 TGAAGTTCAGGTTCTTGTATACGACCCATCTCTGCGGCTTCATCTTTCACACT 3325
Qy 198 -----ValhishepPolysvalleuGlnCysleuanlles 210
Db 3326 TCTCATGTGTCTCTGTTTGACATGCACTTCCCTAAGTCTCTCAGTGTGAATATCA 3385
Qy 210 ertalleuSerGlnlybaryCysGlnuapralatyrproalrglnlilaasapshthetp 230
Db 3386 GCGGCTTAAGTCAAGAAAGGCGAGAGATGTTTACCCGAGACAGATGATGACACATGT 3445
Qy 230 hecysalaglyasplysalaglyarpsaspserscysgln 242
Db 3446 TCTGCGCGGTGACAAAGCAGGTAGAGATCTCTCCAG 3483

RESULT 8

US-09-602-877A-94/c
; Sequence 94, Application US/09602877A
; Patent No. 643707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602.877A
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-94

Alignment Scores:

Pred. No.: 7.15e-69 Length: 735
Score: 725.00 Matches: 130
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.09% Indels: 0
DB: 3 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-602-877A-94 (1-735)

Qy 164 ApeProthrlysaapvalarcpProilaanvalserhisCysProserAlaGlyThr 183
Db 734 CGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATTTCTCTCTGCTGGAGCA 675
Qy 184 LysCysleuvalserglytrpglythrthrlyserproginlilaashepPolysVal 203
Db 674 AAGT 615
Qy 204 LeuGlnCysleuanlleservalLeuSerGlnlybaryCysGlnuapralatyrProalrg 223
Db 614 CTCACAGGCTTGAATATCAGCGTCTAAGTCAAGAAAGGTGCGAGATGCTTACCCGAGA 555
Qy 224 Glnlilaasapshthetphecysalaglyasplysalaglyarpsaspserscysgln 243
Db 554 CAGATTAATGACACCAATGTTCTGCGCGGTCGACAAAGCAGTAAAGACTCTTCCGCAAGGT 495
Qy 244 AspsersglylprovalalCysasnGlyserleuGlnlyleuvalsertrpglyasp 263
Db 494 GATTCTGGGGGGGCTGT 435

Qy 264 TyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCyslysPheThrLys 283
Db 434 TACCCTTGTCCCGGCCCAACAGACCGGGGTCTTACACGAACTCTGCAAGTTCAACCAAG 375
Qy 284 TrpIleGlnIleThrIleGlnAlaAsnSer 293
Db 374 TGGATCCAGGAACCATCCAGGCCAACTCC 345

RESULT 9

US-09-439-313-524
; Sequence 524, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439.313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-439-313-524

Alignment Scores:

Pred. No.: 5.26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 3 Gaps: 6

US-09-936-271b-14 (1-293) x US-09-439-313-524 (1-765)

Qy 1 MetAlaThrAlaArgProPOTrMetTyrValLeuCysAlaLeuIleThrAlaLeu 20
Db 1 ATGCGCCAGACAGAAATCCCTGGGCTGTGTCTCTGGGAGTACTCATC----- 48
Qy 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValserCysAspHisProSerAsn 40
Db 49 CTGGTGTCTC-----GCAGCA 63
Qy 41 ThrValProserGlyserAsnleuAsnleuGlyAlaGlyAlaGlyAlaAsnSer 60
Db 64 TCCTCTCTCTCTCTGT----- 78
Qy 61 AspaAspserserserArgIleIleAsnGlyserAspCysAspMecHisThrGlnProTyr 80
Db 79 -----AGTGCAGGCCAAATCATTAACGCGAGAGATGACGCCGCTGCGAGCCCTGG 132
Qy 81 GlnAlaAlaLeuLeuAsnArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 133 CAGCGCGCACTGGTCATGAA---AACGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
Qy 101 GlnTrpLeuThrThrAlaAlaHisCysArgGlyValValPheArgValaArgLeuGlnHis 120
Db 190 CAGTGGGT 249
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnIleMecPheGlnIleValLysSerIle 140
DB: 121

Db 250 CACAGCTTGGAGCCGACCAAGAGCCAGAGGCCAGATGGTGAAGGCCAGCTTCCGTA 309
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspMetLeuIleuValLeuAsn 160
Db 310 CGGCAACCCAGAGTACACAGACACCTTGTCTGCTAAGCACTTCATGCTCAAGATTGAC 369
Qy 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 370 GAATCCGTCGACAGCTGACACACCATCCGAGCATGATGCTTCCAGTCCCTACC 429
Qy 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrIlySerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGGCTCGTTCTGCTGGGCTGCTGGCAACGCGCAAGAAG----- 483
Qy 201 ProIlyValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlyAspArgCysGluAspAla 220
Db 484 CCAACCGTGTGAGAGCGCGGAACTGTCGTGTGTCTGAGAGGCTGCAATTAAGCTC 543
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspIlyValArgAsp 239
Db 544 TATGACCCGCTGACACACCCAGCATGTTCTGCGCGGAGGAGGCAAGCCAGAGGAC 603
Qy 240 SerCysGlnIlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnIlyLeuVal 259
Db 604 TCTTGAACGAGTACTCTGGGGGCGCCCTGATCTGCAACGGGTACTTGCAGGGCTTTG 663
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
Db 664 TCTTTCGAAAAGCCCGCTGTGGCCAAAGTTGGCGTCCAGAGTCTTACCAACACTCTGC 723
Qy 280 LysPheThrIlyTrpIleGlnIlyThrIleGlnIlyAsn 292
Db 724 AATTCACGTAGTGATAGAGAAACCGTCCAGGCAAGT 762

RESULT 10

US-09-636-215-524.
; Sequence 524, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-636-215-524

Alignment Scores:
Pried. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47

Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6

US-09-936-271B-14 (1-293) x US-09-636-215-524 (1-765)

Qy 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db 1 ARGCCAGAGCAGGAATCCCTGGGGCTGCTTCCGGGTACTCATC-----GCAGCA 48
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 49 CTGGGTGTC-----GCAGCA 63
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValGluAspAlaArgSer 60
Db 64 TCGCTGCTCTGGT----- 78
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 79 -----AGTCGAGCCAAATCATTAACGCGAGACTGCAAGCCGCACTCGAGCCCTGG 132
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnIlyTyrCysGlyAlaValLeuValHisPro 100
Db 133 CAGCGCGCACTGCTCATGGA---AACGAATTGTTCTGCTCGGGGTCTGAGTGCATCCG 189
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgIlyValValPheArgValArgLeuGlnHis 120
Db 190 CAGTGGTCTCTACGCCGACACTGTTCCAAACTCTTACCACTCGGGCTCGGCTG 249
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnIlyMetPheGlnIlyValIlySerIle 140
Db 250 CACAGCTTGAAGCCGACCAAGACGAGGAGCCAGATGTGAGAGCCAGCTTCCGTA 309
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspMetLeuIleIlyLeuAsn 160
Db 310 CGGCAACCCAGAGTACACAGACCTTGTCTGCTAAGCACTTCATCTCAAGTTGAC 369
Qy 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 370 GAATCCGTCGACAGCTGACACCATCCGAGCATGATGCTTCCAGTCCCTACC 429
Qy 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrIlySerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGGCTCGTTCTGCTGGGCTGCTGGCAAGGCAAGATG----- 483
Qy 201 ProIlyValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlyAspArgCysGluAspAla 220
Db 484 CCAACCGTGTGAGAGCGCGGAACTGTCGTGTGTCTGAGAGGCTGCAATTAAGCTC 543
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspIlyValArgAsp 239
Db 544 TATGACCCGCTGACACCCAGCATGTTCTGCGCGGAGGAGGCAAGCCAGAGGAC 603
Qy 240 SerCysGlnIlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnIlyLeuVal 259
Db 604 TCTTGAACGAGTACTCTGGGGGCGCCCTGATCTGCAACGGGTACTTGCAGGGCTTTG 663
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
Db 664 TCTTTCGAAAAGCCCGCTGTGGCCAAAGTTGGCGTCCAGAGTCTTACCAACACTCTGC 723
Qy 280 LysPheThrIlyTrpIleGlnIlyThrIleGlnIlyAsn 292
Db 724 AATTCACGTAGTGATAGAGAAACCGTCCAGGCAAGT 762

RESULT 11

US-09-685-166A-524
; Sequence 524, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Davin C.

```

APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C21
CURRENT APPLICATION NUMBER: US/09/685,166A
NUMBER OF SEQ ID NOS: 898
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 524
LENGTH: 765
TYPE: DNA
ORGANISM: Homo sapien
US-09-685-166A-524

Alignment Scores:
Pred. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
Gaps: 6

US-09-936-271B-14 (1-293) x US-09-685-166A-524 (1-765)
QY 1 MetAlaThrAlaArgProPOTrMetTTrValLeuCyAlaLeuLeuLeuLeu 20
Db 1 ATGGCCACAGCAGGAAATCCCTGGGCTGTTCTGGGAGTACCTCATC----- 48
QY 21 LeuGlyValThrGluHisValLeuAlaHisAspValSerCyAspHisProSerAsn 40
Db 49 CTGGGTCTC-----GCGAGA 63
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAspAlaArgSer 60
Db 64 TCGCTGCTCTGCT----- 78
QY 61 AspAspSerSerSerArgGlyLeuLeuGlySerAspCyAspMetHisThrGlnProTrp 80
Db 79 -----AGCTGCAGCCAAATCATATAACGGCGAGACTGCGCCGCTCGCAGCCCTGG 132
QY 81 GlnAlaAlaLeuLeuLeuAspArgProAsnGlnLeuTrpCyAsnGlyAlaValLeuValHisPro 100
Db 133 CAGCGGCACTGGTCTGATGAA--AACGAATTGTTCTGCTGGCGCTCGGTGTCATCCG 189
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyValPheArgValArgLeuGlyHis 120
Db 190 CAGTGGGTGGTGTAGCGGCACTGTTTCCAGAACTCTTACACATCGGCTGGGCTG 249
QY 121 TyrSerLeuSerProValTyrGlySerGlyGlnGlnMetPheGlnGlyValysSerIle 140
Db 250 CACAGTCTTGGGCGGCAACAAGCCAGGCGGAGCGAGATGGTGGAGGCGCAGCTTCGTA 309
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValLeuAsn 160
Db 310 CGGCAACCCAGAGTACAAAGACCTTGGTCCCTAAACGATCCATGATTAATTGAC 369
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyAspProSer 180
Db 370 GAATCCGTGTCGAGTGTGACACATCCGAGCATACAGCATTCGTTGCGAGTGCCTTACC 429

```

```

QY 181 AlaGlyThrLeuCyAsnLeuValSerGlyTyrGlyThrThrLeuSerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGCTCTGTTTGGCTGGGCTCTCTGCGCAACGCGAAGT----- 463
QY 201 ProLeuValLeuGlnCyAsnLeuHisLeuSerValLeuSerGlnValArgCyAsnAla 220
Db 484 CTTACCGTGTGTGAGTGGCGTGAACGTTGCGGTGTCTGAGAGAGTCTGAGTAACTC 543
QY 221 TyrProGlnGlnLeuAspAspThrMetPheCyAsnLeuGlyAspValArgValAsp 239
Db 544 TATGACCCGCTGTACCAACCCAGCATGTTCTGGCGCGAGCGCAAGCAGAAAGAC 603
QY 240 SerCyAsnGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuVal 259
Db 604 TCTTGCAACGCTACTGTGGGGGCGCCCTGATCTGCAACGAGTACTTCCAGGCGCTTGTG 663
QY 260 SerTrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCy 279
Db 664 TCTTGGAAAAGCCCGGTGTGCGCAAGTTGGCGTGCAGGTGTCTTACACAACTCTGC 723
QY 280 LysPheThrLysTrpIleGlnGlyThrIleGlnAlaAsn 292
Db 724 AAATTCATGATGATAGAAACCGTCCAGCGCAGT 762

RESULT 12
US-09-679-426-524
Sequence 524, Application US/09679426
Patent No. 6759515
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqu
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C20
CURRENT APPLICATION NUMBER: US/09/679,426
NUMBER OF SEQ ID NOS: 895
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 524
LENGTH: 765
TYPE: DNA
ORGANISM: Homo sapien
US-09-679-426-524

Alignment Scores:
Pred. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
Gaps: 6

US-09-936-271B-14 (1-293) x US-09-679-426-524 (1-765)
QY 1 MetAlaThrAlaArgProPOTrMetTTrValLeuCyAlaLeuLeuLeuLeu 20
Db 1 ATGGCCACAGCAGGAAATCCCTGGGCTGTTCTGGGAGTACCTCATC----- 48

```

```

OY      21  LeuLlYalThrgLuhIsvLaLeuLaLaAaAspAvalSerCyAspArlHsProSerAa 40
Db      49  CTTCGGTCTC-----GCAAGA 63
OY      41  ThrValProSerGIysSerAaInAspLeuGIyAlaGIyAlaAspAlaArgSer 60
Db      64  TCCTCGCTCTCTGGT-----78
OY      61  AspaAspSerSerAaGlllelleaAngIysSerAaPcyAspMeHIsThGlnProTTrp 80
Db      79  -----AGTGCAGACCAATCAATAAACGGGAGAGACTGACGCCCACTCGAGCCCTGG 132
OY      81  GlnAlaAlaLeuLeuLeuAArgProAaAngIlnLeuTyCySylYAlaValLeuValHsPro 100
Db      133  CAGCGCGCACTGCGTCATCGAaA---MACGAATTCTTTCGTCGGGGCTCTGTCGATCCG 180
OY      101  GlnTrPLeuLeuThralaAlaHIsCyAArgLysLysValPheAArgValAArgLeuGIyHIs 120
Db      190  CAGTGGGTGCTGTCAGCGCAGACTGTTCCAGAACTTCCTACACCATCGAGGTGGGCTCG 240
OY      121  TyrserLeuSerProValTyrgIusSerGIyGlnGlnMetPheGlnGlyValIysSerIle 140
Db      250  CACAGTCTTAAGCGCCGACCAAGAGCCAGAGGAGCCAAATGATGGAAGGCCAGCCTTCCCTA 300
OY      141  ProHIsProGIyTyrserHIsProGIyHIsSerAspaAspLeuMeLeuIleLysLeuAa 160
Db      310  CGGACACCCAGATACAAACACACCCCTTGCTGCTGCTACAGACTTATGCTCATCAAGTTGAC 360
OY      161  ArgAArgIleArgProThrTyLysAspValAArgProIleAaAvalSerSerHIsCyAsProSer 180
Db      370  GAATCCCGTGTCCGAGTGTGACACCAATCCGAGCATCAGCATTCCTTCCAGATGCCCTTACC 420
OY      181  AlaGIyHIsTyrcysLeuValSerGIyTrPTrGIyThrThLysSerProGlnValHIsPhe 200
Db      430  GCGGGGAACTCTTGCTCCCTGTTCTGCTGGAGTCTGCTGGCAACGGCAGAAATG----- 480
OY      201  ProLysValLeuGlnCyLysLeuAaHIsSerValLeuSerGlnLysAArgSyluAspAla 220
Db      484  CCTACCGTGTGCGAGTGCCTGGAACGTGTGCGTGTCTGTGAAGAAGTCTGCACGTAAAGTGC 540
OY      221  TyrProAArgGlnIleAspaAspHrMetPheCyAAla--GlyAspLysAlaGIyAArgAsp 235
Db      544  TATACACCCGCTGTACACCCACGACATGTTCTGGCGCGGAGGCGCAAGACACAGAAAGAC 600
OY      240  SerCyGlnGlyAspSerGIyGIyProValValCyAaAngIysSerLeuGlnGIyLeuVal 255
Db      604  TCTGTCAACGTTGACTCTGGGGGGGCCCTCGATGTGCAGAGGTACTGGAGGGCTTGTG 660
OY      260  SerTrpGIyAspTyrcySylAaArgProAspaArgProGIyValTyTrThAsnLeuCyS 275
Db      664  TCTTTCGAAAGACCCCGTGTGGCCAAAGTTGGCGTGCAGGTGTGTACACCACTCTGC 720
OY      280  LysPheThrIystrPllleGlnGlnTrIleGlnAlaAa 292
Db      724  AAATTCACTGATGEGATAGAGAAACCGTCCAGGCCAGT 762

RESULT 13
US-09-759-143-524
; Sequence 524, Application US/09759143
; Patent No. 6800746
; GENERAL INFORMATION:
; APPLICANT: Xu, Jlangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Hatlocker, Susan L.
; APPLICANT: Jlang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.

```

```

? APPLICANT: Vedvick, Thomas S.
? APPLICANT: Carter, Darrick
? APPLICANT: Li, Samuel
? APPLICANT: Wang, Aijun
? APPLICANT: Skeiky, Yashir A.W.
? APPLICANT: Hepler, William
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
? FILE REFERENCE: 210121.427C23
? CURRENT APPLICATION NUMBER: US/09/759,143
? NUMBER OF SEQ ID NOS: 934
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 524
? LENGTH: 765
? TYPE: DNA
? ORGANISM: Homo sapien
US-09-759-143-524

Alignment Scores:
Pred. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6

US-09-936-271B-14 (1-293) x US-09-759-143-524 (1-765)

OY 1 MetAlaThAlaArgProProffThProffThValleuCySaAlaLeuIleThAlaLeu 20
    |||||
DB 1 ATGGCCACGACGAAATCTCGGGGCTGTTCCCGGGATCCATC----- 48
    |||||

OY 21 LeuGlYValThGluuHISValLeuAlaAsnAspValSerCyAspRHISProSerAsn 40
    |||||
DB 49 CTGGGTGTC-----GCACGA 63
    |||||

OY 41 ThValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaSer 60
    |||||
DB 64 TCCTCTGCTCGGT----- 78
    |||||

OY 61 AspAspSerSerArgIleIleAsnGlySerAspCyAspMetHISGlnProTrp 80
    |||||
DB 79 -----AGCTGCAGCCAAATCTAAACGCCAGACATCCACCCGCACCTCGACCTCG 132
    |||||

OY 81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTrpCyGlyAlaValLeuValHISPro 100
    |||||
DB 133 CAGGCGGCACTGTCATCGAA---AACCAATGTCTGCTCGGCGCTCGGTGCATCCG 189
    |||||

OY 101 GlnTrpLeuLeuThAlaAlaHISCyAspArgIleValValPheArgValAspLeuGlyHIS 120
    |||||
DB 190 CAGTGGGTGCTGTCACCCGACACACTGTTCCAGAACTCTACACCAATCGGCTGGGCTG 249
    |||||

OY 121 TyrSerLeuSerProValTrpGluSerGlyGlnGlnMetPheGlnGlyValValSerIle 140
    |||||
DB 250 CACAGCTTGAAGGCCGACCAAGACGACGAGACCGAGATGGTGGAGCCAGCCTCTCCGTA 309
    |||||

OY 141 ProHISProGlyTyrSerHISProGlyHISerAsnAspLeuMetLeuIleValLeuAsn 160
    |||||
DB 310 CGGACCCACAGATCAACACGACCTTGTCTCGTAAACGCTCAAGCTCATCAATTAAGTGGAC 369
    |||||

OY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerHISCyProSer 180
    |||||
DB 370 GAATCCGTGTCCAGCTGACACCATCCGAGACATACGATGCTTGGCACTGCCATACC 429
    |||||

OY 181 AlaGlyThrIleCyLeuValSerGlyTrpGlyThrThrIleValSerProGlnValHISPhe 200
    |||||
DB 430 GCGGGAAACTTGGCTCGTTCTGGCTGGGGGTCTGCTGCGAAGCGACGAATG----- 483
    |||||

OY 201 ProValValLeuGlnCyLeuAsnHISerValLeuSerGlnValGlyCyGlnAspAla 220
    |||||
DB 484 CTTACCGTGTGACGTGACCTGAACCTGTGCGGTGGTGTCTAGAGAGCTGTGACGTAACTC 543
    |||||

```

QY 221 TyrProArgGlnIleAspThrMetPheCysAla---GlyAspValAlaGlyArgAsp 239
DB 544 TATACCCGCTGTACACCCAGCATGTTCTGCGCGGAGGAGGACCAAGAC 603
QY 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
DB 604 TCCTGCACACGGTGACTGTGGGGGGCCCTGATCTGCACACGGGTACTTGCACAGGCTTGTG 663
QY 260 SerTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
DB 664 TCCTTCGAAAGAGCCCGCTGTGGCCAGTTGCGCGGTGCTTACACCACTCTGCG 723
QY 280 LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn 292
DB 724 AAATTCAGTGAATGATAGAGAAACCGTCCAGGCGCACT 762

RESULT 14
US-09-651-236-524
Sequence 524, Application US/09651236
Patent No. 6818751
GENERAL INFORMATION:
APPLICANT: Xu, Jiaangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitchell, Jennifer L.
APPLICANT: Heflocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kados, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Rafter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.42718C18
CURRENT APPLICATION NUMBER: US/09/651,236
CURRENT FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 865
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 524
LENGTH: 765
TYPE: DNA
ORGANISM: Homo sapien
US-09-651-236-524

Alignment Scores:
Pred. No.: 5,266-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6

US-09-936-271b-14 (1-293) x US-09-651-236-524 (1-765)

QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
DB 1 ATGGCCACAGAGAAATCCCTGGGGGTCTTCTGGGTACTCTATC-----GCGAGA 48
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn 40
DB 49 CTGTGTGTC-----GCGAGA 63
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyIleAspAlaArgSer 60
DB 64 TCGCTCGTCTCTGCT----- 78

QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 79 -----AGCTGCAGCCAAATCATTAAGGAGAGACTGCGCCGCACTGCGAGCTTGG 132
QY 81 GlnAlaIleLeuLeuAspArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
DB 133 CAGGCGGCACTGGTCATGAAA---AACGAATTTGTTGCTCGGGCGGTCTGTGTGATCCG 169
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgValLeuValPheArgValAlaGlnGlyHis 120
DB 190 CAGTGGGTGTGTGACGCGCACCTGTTCACAAATCTCCACACATCGGCGTGGGCTG 249
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
DB 250 CACAGTTTGAAGCCGACCAAGAGCCAGGAGCCAGATGTGAGGCGCACCTCTCCGTA 309
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
DB 310 CGGACCCACAGATACACAGACCCCTTGTCTGCTACGACTCATGTCTCATAGTTGGAC 369
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
DB 370 GAATCCGTGCCAGTCTGACACCATCCGAGCATGACATGCTTGTGCGAGTCCCTTACC 429
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyTyrThrLysSerProGlnValHisPhe 200
DB 430 GCGGGAACTCTTGCTCTGCTGCTGGGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 483
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
DB 484 CTTACCGTGTGACAGTCCGGAACGTGTGCTGTGTCTGTGAGAGGTCTCAGTAAGCTC 543
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspValAlaGlyArgAsp 239
DB 544 TATGACCCGCTGTACACCCAGCATGTTCTGGCGCGGCGGAGGAGGAGGAGGAGGAGGAG 603
QY 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
DB 604 TCCTGCACACGGTGACTGTGGGGGGCCCTGATCTGCACAGGATCTTGCAGGGCTTGTG 663
QY 260 SerTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
DB 664 TCCTTCGAAAGAGCCCGTGTGGCCAGTTGCGCGGTGCTTACACCACTCTGCG 723
QY 280 LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn 292
DB 724 AAATTCAGTGAATGATAGAGAAACCGTCCAGGCGCACT 762

RESULT 15
US-09-386-642-10
Sequence 10, Application US/09386642
Patent No. 6420157
GENERAL INFORMATION:
APPLICANT: Darrow, Andrew
APPLICANT: Qi, Jenson
APPLICANT: Andrade-Gordon, Patricia
TITLE OF INVENTION: Zymogen Activation System
FILE REFERENCE: ORT-1028
CURRENT APPLICATION NUMBER: US/09/386,642
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 1052
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-10

Alignment Scores:
Pred. No.: 5,236-61 Length: 1052

Score: 654.50 Matches: 129
 Percent Similarity: 62.99% Conservative: 48
 Best Local Similarity: 45.91% Mismatches: 87
 Query Match: 40.70% Indels: 17
 DB: 3 Gaps: 7

US-09-936-271b-14 (1-293) x US-09-386-642-10 (1-1052)

```

QY      19  leuLeuLeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisPro 38
      49  CTCCTGCTGCTGGTGGTGTCAAAATCTACTTTGTCCAGGGGTGTGCTTCCGACTACAG 108
QY      39  SerAsnThrValProSerGlySerAsnGlnAspLeuGlyValGlyValAspAla 58
      109  -----GACGACGACGACGTCGAGCGCGCT-----CTTGCT 141
QY      59  ArgSerAspAspSerSerSerArgIleIleAsnGlySerAspCys--AspMetHisThr 77
      142  GCCCCTTTGATGATGATGACAAAGATCGTTGGGGGCTACAACTGTCTAGAAAGCACTCC 201
QY      78  GlnProTroglnAlaAlaLeuLeuLeuArgProAsnGlnLeuThrCysGlyValValLeu 97
      202  CAGCCCTTGGAGGAGCC--CTGTCGAGAGACGGGCTACTCTGTGGGGGAGCGCTC 258
QY      98  ValHisProGlnIleProLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArg 117
      259  ATGCCCCCGAGATGGCTCCGACAGAGCCCACTGCTCAAGCCCGCTACATAGTTAC 318
QY      118  LeuGlyHisIleSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyVal 137
      319  CTGGGGCAGCACAACCTCCAGAG--GAGAGGGCTGTGACAGACCCCGACAGCCACT 375
QY      138  LysSerIleProHisProGlyTyrSer-----HisProGlyHisSerAsnAsp 153
      376  GAGTCTTCCCTCCACCCCGGCTTCAACAAGACCTCCCAACAAGACACCGCAATGAC 435
QY      154  LeuMetLeuIleLysLeuAsnArgArgIleArgProThrLysAspValArgProIleAsn 173
      436  ATCATGCTGGTGAAGATGGCATGCCAGTCTCATCATCAGGGCTGTGGAGCCCTCAC 495
QY      174  ValSerSerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThr 193
      496  CTCCTCTCAGCTGTCACCTGTCGACCAAGCTGCTCATTTCCGCTGGGGCAGCAGC 555
QY      194  LysSerProGlnValHisPheProLysValLeuGlnCysLeuAsnIleSerValLeuSer 213
      556  TCCAGCCCCCAGTTACGCTGCTCACACCTTGCAGTGCACCAACATCACCATCATTTGAG 615
QY      214  GlnLysArgCysGlnAspAlaTyrProArgGlnIleAspAspThrMetPheCysAlaGly 233
      616  CACCAAGAGTGTAGAAAGCTTACCCCGGACATCACAGACATCATGTTGTGGCAGC 675
QY      234  ---AspLysAlaGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsn 252
      676  GTCGACAGAAAGGGGCAAGACTCTGTCAGGGTGACTCCGGGGGCTCTGTGTGTAAAC 735
QY      253  GlySerLeuGlnGlyLeuValSerTrogIleAspTyrProCysAlaArgProAsnArgPro 272
      736  CAGTCTCTTCAAGGCATTTATCTCTGGGGCCAGAGATCCGTGTGGATCACCCGAAAGCT 795
QY      273  GlyValTyrThrAsnLeuCysLysPheThrLysTropIleGlnGlnIleGlnAlaAsn 292
      796  GGTGTCTACAGAAAGTCTCAATATATGTGACTGGATTCAGAGAGAGATGAAGAACAAT 855
QY      293  Ser 293
      856  TCT 858
  
```

Search completed: February 26, 2005, 21:27:34
 Job time : 211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:48:31 : Search time 634 Seconds

(without alignments)
2738.422 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608
Sequence: 1 MATRRPMMWVLCALITALL.....VTNNLCKTKWIGTIGIANS 293

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.epool/US09936271/runat_23022005_130128_27307/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.csl -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPEIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09936271.0cgn_1_1480 -runat_23022005_130128_27307
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1608 | 100.0 | 1260 | 17 | US-10-172-118-1411 |
| 2 | 1608 | 100.0 | 1260 | 17 | US-10-295-027-515 |
| 3 | 1608 | 100.0 | 1260 | 17 | US-10-173-999-53 |
| 4 | 1608 | 100.0 | 1260 | 17 | US-10-342-887-1411 |
| 5 | 1608 | 100.0 | 1260 | 17 | US-09-739-907-52 |
| 6 | 1608 | 100.0 | 1260 | 17 | US-09-938-671-52 |
| 7 | 1608 | 100.0 | 1260 | 17 | US-09-739-907-37 |
| 8 | 1608 | 100.0 | 1260 | 17 | US-09-938-722-308 |
| 9 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 10 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 11 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 12 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 13 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 14 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 15 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 16 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 17 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 18 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 19 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 20 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 21 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 22 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 23 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 24 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 25 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 26 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 27 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 28 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 29 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 30 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 31 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 32 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 33 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 34 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 35 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 36 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 37 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 38 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 39 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 40 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 41 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 42 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 43 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 44 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |
| 45 | 1608 | 100.0 | 1260 | 17 | US-09-989-723-308 |

ALIGNMENTS

RESULT 1
US-10-172-118-1411
; Sequence 1411, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Lindsey, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 1411
LENGTH: 1260
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM_012427
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1411

Alignment Scores:

| Pred. No.: | 3,296-187 | Length: | 1260 |
|------------------------|-----------|---------------|------|
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 17 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x US-10-172-118-1411 (1-1260)

```
QY 1 MetAlaThrAlaArgProTfPmEctRPValLeuCyAlaLeuIleThraAlaLeu 20
DB 43 ATGGCTACAGCAAGACCCCTGGATGTGGTCTCTGTCTGTATCAGCCTTGCTT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAlaAspHisProSer 40
DB 103 CTGGGGGTACAGAGCATGTTCTGCCCAATGATGTTCTCTGTACACCCCTTAAC 162
QY 41 ThrValProSerGlySerAsnAspLeuGlyValGlyValGlyValAspArgSer 60
DB 163 ACCGTCCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGGAAGACGCCCGTGG 222
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAlaAspMetHisThrGlnProTfP 80
DB 223 GATGACAGCAGCAGCGCATCATCATGATCCGATCGGATATGACACCCAGCCGTGG 282
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB 283 CAGGCGCGCTGTGGTAAAGGCCCAACCACTTACTGCGGGCGGTGTGGTGCATCA 342
QY 101 GlnTfPLeuLeuThraAlaHisCyAsnGlyValValPheArgValArgLeuGlyHis 120
DB 343 CAGTGGCTGTCAAGCGCCGCTCAAGTCAAGAAAGTTTCAAGATCCGTCCTGGCCAC 402
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 403 TACTCCCTGTCAACAGTTTATGATCTGGGACAGATGTTCCAGGGGGTCAATCATC 462
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuLeuIleIleLeuAsn 160
DB 463 CCCCAACCTGGCTACTCCACCTGGCCACTTAAACGACCTCATGCTCAAACTGAAC 522
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyAspProSer 180
DB 523 AGAAGATTGTCTCCCTAAAGTGTCAACCCATCAAGTCTCCCTCAATGTCTCTCT 582
QY 181 AlaGlyThrIleAspLeuValSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
DB 583 GCGGGGCAAAAGTGTGTGTGTGTGTGGGGGCAACCAAGGCCCCCAAGTCACTTC 642
QY 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleArgCyGlnAspAla 220
DB 643 CCTAAGGTCTCCAGTGTGAATATCAAGCTGTAAAGTCAAGAAAGGTGCGAGTGTCT 702
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIleValIleArgAspSer 240
DB 703 TACCCGAGACAGATATACATCAATGTTCTGTGGCGGAGACAAACAGGTATGAAGCTCC 762
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 763 TGCAGAGGTGATTTCTGGGGGGCTGTGTGTGTGAAGGTCTCCCTGAGGAGATCGGTTC 822
QY 261 TfPGLYAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyIle 280
```

```
DB 823 TGGGAGATTACCTTGTGCCCCGCCCAACAGCCGGTGTCTACAGAACTTGCAG 882
QY 281 PheThrIleTfPLeuGlnGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 923
DB 883 TTCACCAAGTGTATCCAGGAACATTCAGGCCCACTCC 921
```

RESULT 2

US-10-295-027-515

Sequence 515, Application US/10295027

Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.

APPLICANT: Gish, Kurt C.

APPLICANT: Glynn, Richard

APPLICANT: Hevezl, Peter A.

APPLICANT: Mack, David H.

APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

FILE REFERENCE: 018501-012500US

CURRENT APPLICATION NUMBER: US/10/295,027

CURRENT FILING DATE: 2002-11-13

PRIOR APPLICATION NUMBER: US 09/663,733

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: US 60/350,666

PRIOR FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/335,394

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: US 60/340,376

PRIOR FILING DATE: 2001-12-14

PRIOR APPLICATION NUMBER: US 60/347,211

PRIOR FILING DATE: 2002-01-08

PRIOR APPLICATION NUMBER: US 60/347,349

PRIOR FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: US 60/355,250

PRIOR FILING DATE: 2002-02-08

PRIOR APPLICATION NUMBER: US 60/356,714

PRIOR FILING DATE: 2002-02-13

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1386

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 515

LENGTH: 1260

TYPE: DNA

ORGANISM: Homo sapiens

US-10-295-027-515

Alignment Scores:

| Pred. No.: | 3,296-187 | Length: | 1260 |
|------------------------|---|---------------|------|
| Score: | 1608.00 <td>Matches:</td> <td>293</td> | Matches: | 293 |
| Percent Similarity: | 100.00% <td>Conservative:</td> <td>0</td> | Conservative: | 0 |
| Best Local Similarity: | 100.00% <td>Mismatches:</td> <td>0</td> | Mismatches: | 0 |
| Query Match: | 100.00% <td>Indels:</td> <td>0</td> | Indels: | 0 |
| DB: | 17 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x US-10-295-027-515 (1-1260)

```
QY 1 MetAlaThrAlaArgProTfPmEctRPValLeuCyAlaLeuIleThraAlaLeu 20
DB 43 ATGGCTACAGCAAGACCCCTGGATGTGGTCTCTGTCTGTATCAGCCTTGCTT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAlaAspHisProSer 40
DB 103 CTGGGGGTACAGAGCATGTTCTGCCCAATGATGTTCTCTGTACACCCCTTAAC 162
```

| | | | |
|--|-----|--|------|
| Oy | 41 | ThrValProSerGlySerIleuGlnAspLeuGlyValGlyValGlyValAspAlaArgSer | 60 |
| Db | 163 | ACCGTGCCTCTTGGAGCAACAGAACCTTGGAGCTGGGGCCGGGGAGACGCCGGGTCC | 2222 |
| Oy | 61 | AspAspSerSerSerArgIleIleLeuGlySerAspCysAspMetHisThrGlnProTyr | 80 |
| Db | 223 | GATGACAGCAGCAGCCGATCATCAATGGATCCGACTGGCATATGGACAACCCAGCCGTGG | 2822 |
| Oy | 81 | GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro | 1000 |
| Db | 283 | CAGGCGCGCCTGTGCTTAAGGCCCAACAGCTTACTGCGGGCCGGGTGGTGCAATCCA | 3422 |
| Oy | 101 | GlnTyrPheLeuLeuThrIleAlaHisCysArgGlyLeuValPheArgValArgLeuGlyHis | 1200 |
| Db | 343 | CAGTGCCTGCTCAACGGCCGCCCACTCAGAGAAAGATTTTCAGAGTCCCTCTCGGCAC | 4020 |
| Oy | 121 | TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIlySerIle | 1400 |
| Db | 403 | TACTCCCTGTCAACGATTATATGAAATCTGGGACAGCAATGTTCCAGGGGGCTCAATTCATC | 4630 |
| Oy | 141 | ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleuLeuAsn | 1600 |
| Db | 463 | CCCCACCCTGGCTACTCCACCCCTGGCCACTCTTAACGAACTCATGCTCAATCAATGAC | 5222 |
| Oy | 161 | ArgArgGlyIleArgProThrIlyAspValAlaArgProIleAsnValSerSerHisCysProSer | 1800 |
| Db | 523 | AGAAAGAAATTCGTCCTCCCTAAAGATGTCAAGCCCATCAACGCTCTCTCATTTGCTCCTCT | 5822 |
| Oy | 181 | AlaGlyThrIlyCysLeuValSerGlyTyrTPGlyThrThrIlySerProGlnValHisPhe | 2000 |
| Db | 583 | GCTGGGACAAAGTGTCTGGTGTCTGGCTGGGGGACAAACAAAGGCCCCCAAGTGCATTC | 6422 |
| Oy | 201 | ProIlyValIleuGlnCysLeuAsnIleSerValIleuSerGlnIlyArgCysGlyAspAla | 2200 |
| Db | 643 | CCTAAGTCTCTCCAGTGCTTGAATATCAGGGCTCTAAGTCAGAAAGGTGCGAGGATGCT | 7020 |
| Oy | 221 | TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyValGlyArgAspSer | 2400 |
| Db | 703 | TACCCGAGACAATATACATACACCAATGTTCTGGCCGGGTGACAAAGCAGGTAAAGACTCC | 7622 |
| Oy | 241 | CysGlnGlyAspSerGlyIlyProValIleCysAsnGlySerIleuGlnIlyLeuValSer | 2600 |
| Db | 763 | TGCCAGGGTGATTTCTGGGGGGCTGTGGTGTGCATAGGCTCCCTGCAGGGAACTGTGTCC | 8222 |
| Oy | 261 | TyrGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlyS | 2800 |
| Db | 823 | TGGGAGAGATTACCTTGTGTCCCGGCCCAACAGACCGGGTGTCTACAGCAACTCTGCAAG | 8822 |
| Oy | 281 | PheThrIlyTyrIleGlnGlnIlyThrIleGlnAlaAsnSer | 293 |
| Db | 883 | TTCAACCAAGTGGATCCAGAGAAACCATCCAGGCCAATCC | 921 |
| RESULT 3 | | | |
| US-10-173-999-53 | | | |
| ; Sequence 53, Application US/10173999 | | | |
| ; Publication No. US20040005563A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Mack, David H. | | | |
| ; APPLICANT: Gish, Kurt C. | | | |
| ; APPLICANT: Eos Biotechnology, Inc. | | | |
| ; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions | | | |
| ; TITLE OF INVENTION: And Methods of Screening for Modulators of Ovarian | | | |
| ; FILE REFERENCE: 018501-002420US | | | |
| ; CURRENT APPLICATION NUMBER: US/10/173, 999 | | | |
| ; CURRENT FILING DATE: 2002-06-17 | | | |
| ; PRIOR APPLICATION NUMBER: US 60/299, 234 | | | |
| ; PRIOR FILING DATE: 2001-06-18 | | | |
| ; PRIOR APPLICATION NUMBER: US 60/315, 287 | | | |
| ; PRIOR FILING DATE: 2001-08-27 | | | |
| ; PRIOR APPLICATION NUMBER: US 60/350, 666 | | | |
| ; PRIOR FILING DATE: 2001-11-13 | | | |
| ; PRIOR APPLICATION NUMBER: US 60/372, 246 | | | |

[illegible]

Db 823 TGGGAGATTACCTTGTGCCCCGACAGACCGGAGTGTCTACAGAACTCTGCAAG 882
Qy 281 PheThrLysTrpIleGlnGlnThrIleGlnIleAsnSer 293
Db 883 TTCACCAAGTGGATCCAGAAACCATCCAGGCCACTCC 921
RESULT 4
US-10-342-887-1411
; Sequence 1411, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Betnarde, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1411
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1411
Alignment Scores:
Pred. No.: 3,29e-187 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-09-936-271b-14 (1-293) x US-10-342-887-1411 (1-1260)
Qy 1 MetAlaThrAlaArgProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
Db 43 ATGGCTACGACGAAACCCCTGGAGTGGGCTGTGCTCTGATCAACCCCTGCTT 102
Qy 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCysAspHisProSerAan 40
Db 103 CTGGGGGTCAAGAGATTTCTCGCCAAATGATGTTTCTGTGACCAACCCCTTAAAC 162
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValGlnAspAlaArgSer 60
Db 163 ACCGTCCTCTGGGAGCAACGAGACTGGAGCTGGGGCGGGGAAGACCCCGGTG 222
Qy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 223 GATGACAGACGACCGCATCATCATGATCCGACTGGCATATGACACCCAGCGGTG 282
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyValValLeuValHisPro 100
Db 283 CAGGCGCGCTGTGTTAAGGCCCAACAGCTTACTGCGGGCGGTGTGTGATCCA 342
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgIleValPheArgValArgLeuGlyHis 120
Db 343 CAGTGGCTGCTACGCGCGCCCACTGACAGAAAGATTTCAGATCGGTCTCGGCCAC 402
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 403 TACTCCCTGTCAACAGTTTATGATCTGGGACAGCATGTTCCAGGGGTCAATCATC 462

Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 463 CCCACCTGGGTACTCCACACCTGGCCACTTAACGACCTCATCTCATCAACTGAAAC 522
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 523 AGAAGAAATTCGTCACATTAAGATGTGACACCCATCAAGCTCTCATTTGTCTCT 582
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrLysSerProGlyValHisPhe 200
Db 583 GCTGGACAAAGT 642
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlnAspAla 220
Db 643 CCTAAGTCTCTCACTGCTTGAATTCAGCGGTCAATGCAAGAAAGGTGAGATGCT 702
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 703 TACCCGACAGATGATGATGACACCATGTTGTGCGCGGTGACAAAGCAGATGAGACTCC 762
Qy 241 CysGlnGlyAspSerGlyValProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 763 TCCAGGGGTGATTCGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 822
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 823 TGGGAGATTACCTCTGTGCGCGCCCAACGAGACCGGGTGTCTACACAACTCTGCAAG 882
Qy 281 PheThrLysTrpIleGlnGlnThrIleGlnIleAsnSer 293
Db 883 TTCACCAAGTGGATCCAGAAACCATCCAGGCCACTCC 921

RESULT 5

US-09-739-907-52
; Sequence 52, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-739-907-52
Alignment Scores:
Pred. No.: 3,73e-187 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-936-271b-14 (1-293) x US-09-739-907-52 (1-1381)
Qy 1 MetAlaThrAlaArgProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20

```

Db      129 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTGCTCTGATCACAGCTTCTT 188
Qy      21  LeuGlyValThrGluHisValLeuAlaAspAspValSerCyAspHisProSerAsn 40
Db      189 CTGGGGGTTCACAGAGCATGTTCTCGCCAAACATATGTTCTCTGTGACCAACCCCTCTAAC 248
Qy      41  ThrValProSerGlySerAsnGlnAspLeuGlyValaGlyValaGlyValaAspAsn 60
Db      249 ACCGTGCTCTCTGGAGACAAACAGAGCTGGAGCTGGGGCCGGGGAAAGAGCCGGGTGG 308
Qy      61  AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
Db      309 GATGACAGCAGCAGCCGATCATATGATTCGACATCCGATATGACACCCAGCCCTGG 368
Qy      81  GlnAlaAlaLeuLeuLeuLeuArgProAsnGlnLeuTrpCyGlyValaValLeuValHisPro 100
Db      369 CAGGCCCGCTGTTGTCTTAAGGCCCAACAGCTCTACGCGGGGGGTGGTGCATCA 428
Qy      101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValaArgLeuGlyHis 120
Db      429 CAGTGGCTGCTCACGGCCGCTGCTGACAGAAAGTTTTCAGAGTCCGTTCTGGGCAC 488
Qy      121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db      489 TACTCCCTGTCCACAGTTTATGATCTGGGACAGATGTTTCCAGGGGGTCAATCATC 548
Qy      141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValLeuAsn 160
Db      549 CCCACCCCTGCTACTCCACCCCTGGCCACTCTAACACCTCATGCTCAATCAACTGAC 608
Qy      161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCyProSer 180
Db      609 AGAAGATTGTCTCCACTTAAGATGACAGACCCATCAACGTTCTCTCATGTGCTCTCT 668
Qy      181 AlGlyThrLysCyLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db      669 GCTGGGACAAAGCTGTGTGTGTGTGTGGGGGACAAACAAAGCCCCCAAGTGCATTC 728
Qy      201 ProLysValLeuGlnCyLeuAsnIleSerValLeuSerGlnLysArgCyGlyAspAla 220
Db      729 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGTCTTAAGTCAAGAAAGTCCGAGATGCT 788
Qy      221 TyrProArgGlnIleAspAspThrMetPheCyAspAlaGlyAspLysArgIleAspSer 240
Db      789 TACCCGACAGACAGATGACACCATATCTTGTCCCGGTACAAAGACGATGAGACTCC 848
Qy      241 CyGlnGlyAspSerGlyGlyProValValCyAspGlySerLeuGlnGlyLeuValSer 260
Db      849 TGCACAGGTGATCTGGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 908
Qy      261 TyrGlyAspTyrProCyAspAlaArgProAsnArgProGlyValTyrThrAsnLeuCyAsp 280
Db      909 TGGGGAGATTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 968
Qy      281 PheThrLysThrIleGlnGlnIleGlnIleGlnAlaAspSer 293
Db      969 TTCAACAAAGTGAATCCAGAAACCATCCAGGCCCAACTCC 1007

```

RESULT 6

```

US-09-938-671-52
; Sequence 52, Application US/09938671
; Publication No. US2004002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/938,671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692

```

```

; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-52

Alignment Scores:
Pred. No.: 3,73e-187 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271B-14 (1-293) x US-09-938-671-52 (1-1381)
Qy      1  MetAlaThrAlaArgProTrpMetCysValLeuIleThrAlaLeu 20
Db      129 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTGCTCTGATCACAGCTTCTT 188
Qy      21  LeuGlyValThrGluHisValLeuAlaAspAspValSerCyAspHisProSerAsn 40
Db      189 CTGGGGGTTCACAGAGCATGTTCTCGCCAAACATATGTTCTCTGTGACCAACCCCTCTAAC 248
Qy      41  ThrValProSerGlySerAsnGlnAspLeuGlyValaGlyValaGlyValaAspAsn 60
Db      249 ACCGTGCTCTCTGGAGACAAACAGAGCTGGAGCTGGGGCCGGGGAAAGAGCCGGGTGG 308
Qy      61  AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
Db      249 ACCGTGCTCTCTGGAGACAAACAGAGCTGGAGCTGGGGCCGGGGAAAGAGCCGGGTGG 308
Qy      309 GATGACAGCAGCAGCCGATCATATGATTCGACATCCGATATGACACCCAGCCCTGG 368
Db      369 CAGGCCCGCTGTTGTCTTAAGGCCCAACAGCTCTACTCTGCGGGGTGTGTGTGTGTGT 428
Qy      81  GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCyGlyValaValLeuValHisPro 100
Db      369 CAGGCCCGCTGTTGTCTTAAGGCCCAACAGCTCTACTCTGCGGGGTGTGTGTGTGTGT 428
Qy      101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValaArgLeuGlyHis 120
Db      429 CAGTGGCTGCTCACGGCCGCTGCTGACAGAAAGTTTTCAGAGTCCGTTCTGGGCAC 488
Qy      121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db      489 TACTCCCTGTCCACAGTTTATGATCTGGGACAGATGTTTCCAGGGGGTCAATCATC 548
Qy      141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValLeuAsn 160
Db      549 CCCACCCCTGCTACTCCACCCCTGGCCACTCTAACACCTCATGCTCAATCAACTGAC 608
Qy      161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCyProSer 180
Db      609 AGAAGATTGTCTCCACTTAAGATGACAGACCCATCAACGTTCTCTCATGTGCTCTCT 668
Qy      181 AlGlyThrLysCyLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db      669 GCTGGACAAAGCTGTGTGTGTGTGTGGGGGACAAACAAAGCCCCCAAGTGCATTC 728
Qy      201 ProLysValLeuGlnCyLeuAsnIleSerValLeuSerGlnLysArgCyGlyAspAla 220
Db      729 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGTCTTAAGTCAAGAAAGTGGCGAGTGTCT 788
Qy      221 TyrProArgGlnIleAspAspThrMetPheCyAspAlaGlyAspLysArgIleAspSer 240
Db      789 TACCCGACAGACAGATGACACCATATCTTGTCCCGGTACAAAGACGATGAGACTCC 848
Qy      241 CyGlnGlyAspSerGlyGlyProValValCyAspGlySerLeuGlnGlyLeuValSer 260

```

```
Db      849 TCCAGGAGTATTCGGGGGCGCTGTGCTGCATGCTCCCTGCAGGAGACTCGTGTCC 908
Qy      261 TTPGLVAspTYrProCYsa1AaRgProAaAaRgProG1VAlYrThraSnuLeuCYslys 280
Db      909 TGGGGAGATTACCTTGTGTGCGCGGCCCAACAGACCGGGTGTCTTACAGAACTCTGCAG 968
Qy      281 PheThrLYeTrp1LeGlnGluThrl1eGln1AaAaSer 293
Db      969 TTCACCAAGTGGATCCAGGAACCAATCCAGGCCCACTCC 1007

RESULT 7
US-09-739-907-37
; Sequence 37, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739, 907
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348, 457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070, 567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,c,g, or c
US-09-739-907-37

Alignment Scores:
Pred. No.: 4,24e-187 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 9

US-09-936-271B-14 (1-293) x US-09-739-907-37 (1-1516)
Qy      1 MetAlaThra1AaRgProRTrpMetTTPVAlLeuCYsa1AaLeu1leThra1AaLeuLeu 20
Db      254 ATGGCTACAGCAAGACCCCTGGAGTGGTGTCTGTGCTGTGTCACAGCTTGCTT 313
Qy      21 LeuG1VAlThrG1uH1sVAl1eua1AaAaAaSPVAlSeRcYsaSPH1sProSeAaAa 40
Db      314 CTGGGGGTCAAGAGATGTTCTCGCCAACATGATGTTTCCGTGACCAACCCCTTAAC 373
Qy      41 ThrValProSeRc1YSeRaNgn1AaPLeuG1VAlAG1VAlG1uAaSP1AaRgSeR 60
Db      374 ACCGTGCTCTGGGAGCAACGAGACTGGGAGCTGGGGCGGGGAAGACGCCGCTCG 433
Qy      61 AaPaSPSeRSeRrG1le1leAaNG1YSeRaNcYsaSPMeH1sThrGInProTrp 80
Db      434 GATGACAGAGCAGCGCATCATCATGATCCGACTGCAATATGACACCAAGCCGCTGG 493
Qy      81 GlnAlAAlaLeuLeuAaRgProAaNG1nLeuTYrCYsG1VAlAValLeuValH1sPro 100
Db      494 CAGGCGGCGCTGTGCTTAAGGCCCAACAGCTTACTGCGGGGGGATGTGGTGATCCA 553
Qy      101 GlnTrpLeuLeuThra1AaH1sCYsARgLYsVAlPheARgVAlARgLeuG1YH1s 120
```

```
Db      554 CAGTGGCTGCTCACCGCGCCCACTGCAGGAAGAAATTTCAGAGTCCGTCCGCGCAC 613
Qy      121 TYrSeRLeuSeRProVAlTYrG1uSeRc1YgInG1nMetPheGlnG1VAl1YsSeR1le 140
Db      614 TACTCTGTACACAGTTTAATCTGGGACAGAGATGTTCCAGGGGGTCAAAATCCATC 673
Qy      141 ProH1sProG1YTYrSeRH1sProG1YH1sSeR1AaAaPLeuMeTLeu1leYsLeuAa 160
Db      674 CCCGACCTGGGTACTCCCACTGGCCACTTAAACCACTTAACTCATGATACTGAAC 733
Qy      161 ArgATg1leAaRgProThrLYsAaPValAaRPro1leAaAaVAlSeRSeRH1sCYsProSeR 180
Db      724 AGAAGAAATTCGTCACCAATTAAGATGTCAAGACCAATCAACGTTCTCTCATTTGTCCTCT 793
Qy      181 AlaG1YThrLYsCYsLeuVAlSeRc1YTrpG1YThrThrLYsSeRProG1VAlH1sPhe 200
Db      794 GCTGGGACAAAGTGTGTGTGTGTGGGGGACCAACCAAGAGCCCAAGTGCACATTC 853
Qy      201 ProLYsVAl1eUGInCYsLeuAaH1SeSeRVal1eUSeRc1nLYsARCYsG1uAaSP1A 220
Db      854 CCTAAGTCTCTCACTGCTTGAATTCAGCGTCAATCAAGAAAGGTGGAGGATCTCT 913
Qy      221 TYrProARgGln1leAaSPThrMeCPheCYsa1AG1YsAPLYsAlG1YARgAaSPSeR 240
Db      914 TACCCGAGACAGATGATGACACCATGTTCTGCGCGGTGCAAGCAGTAGAGACTCC 973
Qy      241 CYsGInG1YsAPSeRc1YG1YProVAlVAlCYsAaNG1YsAR1eUGInG1YLeuVAlSeR 260
Db      974 TCCAGGGGTATTCGGGGGGCTGTGTGTGTGCAATGAGCTCCCTGCAGGAGACTCGTGTCC 1033
Qy      261 TTPGLVAspTYrProCYsa1AaRgProAaAaRgProG1VAlYrThraSnuLeuCYslys 280
Db      1034 TGGGAGATTACCTTGTGTGCGCGGCCCAACAGACCGGGTGTCTAACAAGACTCTGCAAG 1093
Qy      281 PheThrLYeTrp1LeGlnGluThrl1eGln1AaAaSer 293
Db      1094 TTCACCAAGTGGATCCAGGAACCAATCCAGGCCCACTCC 1132

RESULT 8
US-09-936-671-37
; Sequence 37, Application US/09936671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/936, 671
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348, 457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070, 567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070, 658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,c,g, or c
US-09-936-671-37

Alignment Scores:
Pred. No.: 4,24e-187 Length: 1516
```


Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-936-671-37 (1-1516)

QY 1 MetAlaThrAlaArgProPrtPrtMetTyrValLeuCyAlaLeuLeuLeuLeu 20
DB 254 ATGGCTTACAGAACACCCCTGATGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 313
QY 21 LeuGlyValThrGluIleValLeuAlaAspAspValSerCyAspPheIleProSerAsn 40
DB 314 CTGGGGGTTCACAGACATGTTCTGTCCCAACATATGTTCTGTGTGACACACCTCTTAC 373
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAlaAspValSer 60
DB 374 ACCGTGCTCTGTGGAGCAACACAGACCTGGGAGCTGGGGCCGGGAGAACCCGGTTCG 433
QY 61 AspAspSerSerSerAlaGlyIleIleAsnGlySerAspCyAspMetHleThrGlnPrtP 80
DB 434 GATGACAGACAGACACCCATCATCATCAATGATCCGACTGCGATATGACACACCCGCTGG 493
QY 81 GlnAlaAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 494 CAGGCGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 553
QY 101 GlnThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 554 CAGGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 613
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLeuSerIle 140
DB 614 TACTCCCTGTACACAGTTTATGATCTGGGACAGATGTTCCAGGGGGTCAATTCATC 673
QY 141 ProHleProGlyTyrSerHleIleProGlyHleSerAspAspMetLeuIleValLeuAsn 160
DB 674 CCCACCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 733
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHleIleValLeu 180
DB 734 AGAAGATTCGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 793
QY 181 AlIGlyThrIleGlySerLeuValSerGlyTyrGlyThrThrIleSerProGlnValHlePhe 200
DB 794 GCTGGGCAAAAGTCTGT 853
QY 201 ProIleValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnIleValArgCyAlaAspAla 220
DB 854 CTTAAGGTCTCTCAAGTCTGTGAATATACGCTGTAAAGTCAAGAAAGTGTGGAGATGCT 913
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAspValArgAspSer 240
DB 914 TACCCAGACAGATAGATGACACCATTTCTGCCGCCGCTGTCAAAAGACGATAGACTCC 973
QY 241 CyeGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyValLeuValSer 260
DB 974 TGCAGAGGTATCTGGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033
QY 261 TrpGlyAspTyrProCyAlaArgProAspAspArgProGlyValTyrThrAsnLeuCyVal 280
DB 1034 TGGGAGATTAACCTTGT 1093
QY 281 PheThrIleTyrTrpIleGlnGlnIleGlnAlaAsnSer 293
DB 1094 TTCACCAAGTGTGATCCAGAAACCATCATCAGGCCCAACTCC 1132

RESULT 9
US-09-989-722-308
Sequence 308, Application US/09989722
Patent No. US2002072067A1
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P27301C63
CURRENT APPLICATION NUMBER: US/09/989,722
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089588
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 4,45e-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x US-09-989-722-308 (1-1570)

QY 1 MetAlaThrAlaArgProTfPmETrPValLeuCyAlaLeuLeuThzAlaLeuLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGCTTATATACAGCCCTTGCTT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 414 CTGGGGGTACAGACATGTTTCTGCCAACAATGATGTTTCTGTGACCAACCCCTTAAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB 474 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGGAAGACCCCGGTGG 533

Qy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80
Db 534 GATGACGACGACGACCCCATCATATGATCCAGTCCATATGACACCCACCCGCTGG 593
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 594 CAGGCGCGCTGTTGCTAAAGGCCCAACAGCTCTACGCGGGGGGGTGGTGGTCACTCA 653
Qy 101 GlnTyrLeuLeuTyrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGCTGCTCAACGCGCCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCGAC 713
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerTle 140
Db 714 TACTCCCTGTGACAGTTTATGAAATCTGGGGAGCAGATGTTCCAGGGGGGTCAAAATTCATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuValAsn 160
Db 774 CCCCACCCCTGGCTACTCCACCCCTGGCCACTCTAACACCTCATGCTCACTAACTGAC 833
Qy 161 ArgArgLysArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGATTCGTCCCACTAAAGATGTCAGACCCATCAAGCTCTCTCATTTGCCCTCT 893
Qy 181 AlIGlyThrLysCysLeuValSerGlyTyrGlyTyrThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGCAAGAGCTGTGTGTCTGGCTGGGGCAACCAAGAGCCCCCAAGTGCATCTTC 953
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlyAspAla 220
Db 954 CTTAAGGTCTCTCCAGTCTTGAATATACGCTGTCTAAGTCAAGAAAGCTCGAGATGCT 1013
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGACAGATAGATGACCATGTTCTGCGCCCGTCAACAAAGCAGGTAGAGACTCC 1073
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValLysSer 260
Db 1074 TGCAGGGGTATCTGGGGGGGCGTGTGTCTGCATAGCTCCCTCGCAGGGAATCTGTGCTCC 1133
Qy 261 TyrGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGAGATTAACCTTGTGCGCGCCCAACAGACGCGGGTGTCTAACGAACTCTTGCAAG 1193
Qy 281 PheThrLysTyrIleGlnGlnTyrIleGlnAlaAsnSer 293
Db 1194 TTCACCAAGTGAATCCAGAAACCATCAGGCGCAACTCC 1232

RESULT 10
US-09-989-723-308
Sequence 308, Application US/09989723
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlt, Steven, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavyn, Ivar J.
APPLICANT: Napiet, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C62
CURRENT APPLICATION NUMBER: US/09/989,723
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

```
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088858
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088861
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088876
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/089105
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/089440
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089512
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089514
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089532
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089538
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089598
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089599
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089600
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089653
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089907
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089908
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089947
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089948
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089952
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/090246
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090252
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090254
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090349
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090355
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090429
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090431
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090435
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090444
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090445
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090472
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090535
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090540
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090542
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090676
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090678
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090690
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090862
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090863
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/091360
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091478
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091544
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 4,45e-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 9 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-989-723-308 (1-1570)
QY 1 MetalaThraAlaArgProProTfPmEtrPValLeuCyAlaLeuIeThraAlaLeu 20
Db 354 ATGGCTACAGCAAGACCCCTCGATGTGGTCTGTGTCTGTATCAAGCCTTGGT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
Db 414 CTGGGGGTCAAGACAGATGTTCTCGCAATGATGTTTCTGTGACCAACCTCTAAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValGlyValAspAlaArgSer 60
Db 474 ACCGTGCCCTCTGGAGCAACCGAGCTGGAGCTGGAGCCGGGAGAGAGCCCGGTG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80
Db 534 CATGACACGACGACCGCATCATCATGATCCAGCTCGATATCAACACCGCGGTGG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 594 CAGGCGCGGTGTCTTAAGGCCCAACGACTCTACTCGGGGCGGTGTGTGTGATCCA 653
QY 101 GlnTyrPleuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGCTGCTTACCGCCGCGCATCGAGAGAAAGTTTTCAGAGTCCGTCTGAGCAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValYssSerIle 140
Db 714 TACTCCCTGTCAACGATTATGCAATCTGGGACGAGATGTTCCAGGGGTCAATTCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleYsLeuAsn 160
```

Db 774 CCCACCCCTGGCTACTCCACCTGGCCACTCTAAGCAGCTCATGCTCATCAAGTGAAC 833
Qy 161 ArgArgIleArgProThrIleValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGAAATTCCTCCACTCAAGATGACAGCCATCAACGCTCTCATTTGCTCTCT 893
Qy 181 AlagIlyThrIleCysLeuValSerGlyTTrpGlyThrThrIleSerProGlnIValHisPhe 200
Db 894 GCTGGGACAAAGTCTTGCTGTCTGGCTGGGGGACACCAAGAGCCCCCAAGTCACTTC 953
Qy 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
Db 954 CCTAAGGCTCTCCAGTCTTGAATATCAAGCTGTCAAGTGAAGAAAGTGGAGAGATGCT 1013
Qy 221 TyTrpArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValArgIleValArgSer 240
Db 1014 TACCCGAGACAGATGATGACACCATGTTCTGCCCCGTGCAAGCAGGTGAGACTCC 1073
Qy 241 CysGlnIleAspSerGlyIleProValValCysAsnIleSerLeuGlnIleValSer 260
Db 1074 TGCGAGGCTGATTTGGGGGGCTGTGTCTGCAATGCTCTCTGCGAGGACTGTCTTC 1133
Qy 261 TrpIleAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIle 280
Db 1134 TGCGAGATTAACCTTGCTGGCCGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1193
Qy 281 PheThrIleTyrPileGlnIleThrIleGlnIleAlaIleAsnSer 293
Db 1194 TTGACCAAGTGATCCAGAAACCATCCAGGCCAATCC 1232

RESULT 11
US-09-989-279-308
Sequence 308, Application US/09989279
Patent No. US20072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Betsstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlesen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2730P1C56
CURRENT APPLICATION NUMBER: US/09/989,279
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544

PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

| pred. No.: | 4 456-187 | length: | 1570 |
|------------------------|-----------|---------------|------|
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x US-09-989-279-308 (1-1570)

1 MetalathrAlaArgProTfPmetTfValLeuCyAlaLeuLeuThraLeuLeu 20
354 ATGGCTACGAAAGACCCCTGATGTGGGTCTGTGCTGATCACACCTTGGCTT 413

21 LeuGlyValThrGlnHisValLeuAlaMetAlaMetValSerCyAspHisProSerMet 40
414 CTGGGGGCTCACAGACATGTTCTCCCAACATGATGTTCTTGACCAACCCCTTAC 473

41 ThrValProSerGlySerMetGlnMetLeuGlyAlaGlyAlaGlyAlaArgSer 60
474 ACCGTGCTCTTGGAGACACCAAGACCTTGGAGCTGGGGCCGGGAAAGCCGGCTG 533

61 AspAspSerSerSerArgIleIleAlaGlySerMetCyAspMetHisThrGlnProTfP 80
534 GATGACAGACAGACCGCATCATCATGATGATCCGATGATATGATCACACCCAGCGTGG 593

81 GlnAlaAlaLeuLeuLeuArgProMetGlnLeuTyrCyGlyAlaValLeuValHisPro 100
594 CAGGCCGCGCTGTCTTAAGGCCCAACGACTTCTGCGGGCCGTGTGTGATCAACA 653

101 GlnTfPLeuLeuThraAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
654 CAGTGGCTCTCACCGCCGCCCATCTGCAGAAAGATTTCAGATCCGTCTCGGCCAC 713

121 TyrSerLeuSerProValTyrGlnSerGlyGlnMetPheGlnGlyValIleSerIle 140
714 TACTCCCTGTACCAAGTTATGATATCTGGGACAGATGATCTCAGGGGGTCAATTCATC 773

141 ProHisProGlyTyrSerHisProGlyHisSerMetAlaMetLeuLeuIleValSerMet 160
774 CCCCACTGCTACTCCACCTGCGCACTTAAAGACATGATCTCAATCAACTGAAC 833

161 ArgArgIleArgProThrTyrAspValArgProIleAlaMetValSerSerHisCyProSer 180
834 AAGAGATTCTGCCACCTTAAGATGTCAAGCCCATCAACGTCTCTCATTTGCTCCCTT 893

181 AlaGlyThrValCyLeuValSerGlyTyrGlyThrThrValSerProGlnValHisPhe 200
894 GCTGGGACAAAGTGTGTGTGTCTGGCTGGGGACAAACCAAGGCCCAAGTGCACCTTC 953

201 ProIleValLeuGlnCyLeuValMetIleSerValLeuSerGlnValArgCyGlyValPhe 220
954 CCTAAGTCTCTCAGAGCTTGAATATCAGCGCTTAAGTCAAGAAAAGGTGACGAGTGTCT 1013

221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAlaGlyArgAspSer 240
1014 TACCCGAGACAGATGATGACCAATGTTCTGCGGGGACAAAGGACAGATGACGCTCC 1073

241 CyGlnGlyAspSerGlyGlyProValValCyAspMetGlySerLeuGlnIleValSer 260

Db 1074 TGGCAGGATGTTTGGGGGGCCCTGTGTCGATGGCTCCCTGCAGGACTGTGTCC 1133
Qy 261 TTPGLApyTyProCyaaIaXpPcaBnArgrProGlyValTyThrAenLeuCyals 280
Db 1134 TGGGAGATTAACCTTGCTGGCCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAG 1193
Qy 281 PheThrIyTTPILeGInGluThrIleGlnIaBnser 293
Db 1194 TTCACCAAGTGATTCAGAAACATCCAGGCCCAACTCC 1232

RESULT 12
US-09-989-727-308
Sequence 308. Application US/09989727
Patent No. US20020072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bostein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Bacon, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlesen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1G65
CURRENT APPLICATION NUMBER: US/09/989,727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089633
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948

;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090662
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 4,45e-187
Score: 1608.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%

Length: 1570
Matches: 293
Conservative: 0
Mismatches: 0

Query Match: 100.00%
DB: 9
Indels: 0
Gaps: 0
US-09-936-271b-14 (1-293) x US-09-989-727-308 (1-1570)

QY 1 MetalaarhAlaArpProPrtPmeTrrPValLeuCyAlaLeu1eThra1aLeuLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGGGGCTCTGGCTCTGATCCACACCTTGGCTT 413
QY 21 LeuGlyValThrhGluHlValLeuAlaLeuAlaPValSerCyAspHisProSerAn 40
DB 414 CTGGGGGTACAGACATGTTCTCCCAACATGATGTTCTGTGACCACTCTTAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSer 60
DB 474 ACCGTGCTCTCGGAGCAACAGACTGGAGCTGGGGCCGGGAGAACCCCGGTG 533
QY 61 AspAspSerSerSerArgLleLeuAnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 534 GATGACAGCAGCAGCCGATCATCATGATGATCCGATGATATGACACCCAGCGGTG 593
QY 81 GlnAlaAlaLeuLeuAlaArgProAsnGlnLeuTyroGlyAlaValLeuValHisPro 100
DB 594 CAGGCGCGCTGTCTTAAGCCCAACAGCTTACTGCGGGCCGTTGGATGATCCA 653
QY 101 GlnTrpLeuLeuThraAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
DB 654 CAGTGGCTGTCAACGCCGCCCACTGAGGAAGAAATTTTCAAGTCCGTCTGGGCAC 713
QY 121 TyroSerLeuSerProValTyroGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
DB 714 TACTCCCTGTCAACAGTTTATGATGATGAGGAGCAAGTGTCCAGGGGGTCAATTCATTC 773
QY 141 ProHisProGlyTyroSerHisProGlyHisSerAsnAspMetLeuLleValSerLeuAn 160
DB 774 CCCACCTGCTACTCCACCTGACCTTAACGACTATGCTCATCAACTGAAAC 833
QY 161 ArgArgLleArgProThraLysAspValArgProLleAsnValSerSerHisCyAspProSer 180
DB 834 AAGAGAAATTCGTCACATTAAGATGTCAGACCACTCAAGTCTCTCATTTGTCCTCT 893
QY 181 AlaGlyThrLysCyLeuValSerGlyTyroGlyThrLysSerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGTGTGGTGTCTGGCTGGGGGAGCAACCAAGAGCCCCAAGTGCATTC 953
QY 201 ProLysValLeuGlnCyLeuAsnLleSerValLeuSerGlnLysArgCyLeuAspAla 220
DB 954 CCTAAGTCTCTCACTGCTTGAATATCAGCGTCAAGTCAGAAAAGTGCAGAGATGCT 1013
QY 221 TyrProArgGlnLleAspAspThrMetPheCyAlaGlyAspLysAlaGlyArgAspSer 240
DB 1014 TACCCGAGACAGATGATGATCAACCATGTTCTGCGCGGAGCAAGAGGTGAGACTTC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TCCCAAGGTGATTCGGGGGGCTGTGGTGTGCATATGCTCTCAGAGGACTCGTGTCC 1133
QY 261 TrpGlyAspTyroProCyAlaArgProAsnAspProGlyValTyroThrAsnLeuCyLys 280
DB 1134 TGGGAGATTACCTTGTGCCCCGCAACAGCGGGGTGTACACGAACCTTGCAAG 1193
QY 281 PheThrLysTrpLleGlnGlnThrLleGlnAlaAsnSer 293
DB 1194 TTCACCAAGTGCATCCAGGAACCATCCAGGCAACTCC 1232

RESULT 13

US-09-989-731-308
; Sequence 308, Application US/09989731

; Patent No. US20020103125A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David

APPLICANT: Deenoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Peoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24

```
/ PRIOR APPLICATION NUMBER: 60/090444
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090445
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090472
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090535
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090540
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090542
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090676
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090678
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090690
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090682
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090863
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/091360
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091478
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091544
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09
```

Alignment Scores:

```
Pred. No.: 4,45e-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
```

US-09-936-271B-14 (1-293) x US-09-989-731-308 (1-1570)

```
QY 1 MetAlaThrAlaArgProPOTPMetTPPValLeuCYsAlaLeuIleThrAlaLeuLeu 20
Db 354 ATGGCTACAGCAGACCCCTGGATGTGGTCTCTGTCTGTATCAGCCTTGCTT 413
QY 21 LeuGlyValThrLuhIstValLeuAlaAenAenAspValSerCyAspHisProSerAen 40
Db 414 CTGGGGGTCAACAAGATGTTCTGGCCAAATGATGTTCTGTGACCACTCCCTTAAC 473
QY 41 ThrValProSerGlySerAenGlnAspLeuGlyValAGlyValAGlyGluAspAlaA9Ser 60
Db 474 ACCGTCCCTCTGGAGCAACCAAGACCTGGAGCTGGGGCCGGGAAAGACGCCCGGTG 533
QY 61 AspAspSerSerSerArgIleIleAenGlySerAspCyAspMetHisThrGlnProTTP 80
```

```
Db 534 GATGACAGCAGCAGCCGCTCATCATGATCCGACTGGCATATGACACACCACCGCTG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAenGlnLeuValTyCGlyValAlaLeuValHisPro 100
Db 594 CAGCCGCGCTGTGTGTAAGCCCAACAGCTCTACTGGGGGGGTGTGGGTGACATCA 653
QY 101 GlnTPLeuLeuThrAlaAlaHisCyAspArgLysValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGCTCTCACCGCCGCCCATCGACAGGAAGAAAGTTTCAAGATCCGTCTCGGCCAC 713
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 714 TACTCCCTGTCAACAGTTATGATCTGGGCGACGATGTTCCAGGGGTCAAAATCCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAenAspLeuMetLeuIleLysLeuAen 160
Db 774 CCCACCCCTGGCTACTCCACCCCTGGCCACTTAACGACTATGCTCATCAACTGAAAC 833
QY 161 ArgArgIleArgProThrLysAspValArgProIleAenValSerSerHisCyAspProSer 180
Db 834 ABAAGATTCTGCTCCACTTAAGATGTCAGACCCCATCAAGTCTCTCATTTGCTCCT 893
QY 181 AlaGlyThrLysCyAspLeuValSerGlyTyrGlyThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGACAAAGTGTGGTGTCTGGCTGGGGGACCAACCAAGAGCCCAAGTGCATTC 953
QY 201 ProLysValLeuGlnCyAspLeuAenIleSerValLeuSerGlnLysArgCyAspGluAspAla 220
Db 954 CCTAAGTCTCTCAGGCTTGTAATACAGCGGTCAAGCAAAAGGTCAGAGATGCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAspAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGAGACATATGATGATGACACCATGTTCTGCGCGGGACAAAGGACGTAGAGCTCC 1073
QY 241 CySGlnGlyAspSerGlyGlyProValValCyAspGlnLysSerLeuGlnGlyValSer 260
Db 1074 TCCCAAGGTGATTCGGGGGGCTGTGTCTCAATGCTCCCTCAGGAGCATCGTGTCC 1133
QY 261 TPpGlyAspTyrProCyAspAlaArgProAenAspProGlyValTyrThrAenLeuCyLys 280
Db 1134 TGGGAGATTACCTTGTCGCGGCCCAACACAGCGGGTGTCTACGAACTCTGCAAG 1193
QY 281 PheThrLysTPPLeuGlnGlnIleThrIleGlnAlaAenSer 293
Db 1194 TTCACCAAGTGTGATCCAGGAAACCATCCAGGCCAATCTCC 1232
```

RESULT 14
US-09-989-732-308
Sequence 308, Application US/09989732
Patent No. US20020123463A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Denoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerder, Hanspeter
APPLICANT: Gerlitsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OR INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C57
CURRENT APPLICATION NUMBER: US/09/989,732
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690

/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090697
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090698
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090699
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090700
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 4,456-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-989-732-308 (1-1570)

QY 1 MetAlaThrAlaArgProPOTTPMettPvAlleuCyAlaLeuIleThrAlaLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGAGTGGGTCTGTGCTGTGCTGTGCTGTGCT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSer 40
DB 414 CTGGGGGTCAAGAGCATGTTCTGGCAACAATGATGTTCTGTGACCACTCTTAAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGlnAspAlaArgSer 60
DB 474 ACCGTGCTCTGTGGAGCAACCAAGACCTGGAGCTGGGGCCGGGGAAGCCGCGTGG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTTP 80
DB 534 GAAGACAGACAGCGCGCATCATCAATGATCCGACTGGCATATGACACCCGCGTGG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB 594 CAGGCGCGCTGTGCTTAAGGCCCAACCACTTACTGCGGGCGGTGTGGTCAATCA 653
QY 101 GlnTTPLeuLeuThrAlaAlaHisCyArgGlyValAlaPheArgValArgLeuGlyHis 120
DB 654 CAGTGGCTGTCAAGCGCGCCACTGCAAGAAAGATTTCAGAGTCCGTCTCGGCAC 713
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValHisSerIle 140
DB 714 TACTCCCTGTCAACCAATTAATGATCTGGGCAACAATGTTCCAGGGGGTCAATCAATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValHisLeu 160
DB 774 CCCACCTGGCTACTCCCACTGGGCACTTAAGACCTCAATGCTCAATCAATGAAAC 833
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyProSer 180

DB 834 AAGAAGATTCCGCCACTAAAGATGTCAGACCCATCAAGTCTCTCATTTGCTCCT 893
QY 181 AlaGlyThrIleValSerGlyTyrGlyThrThrIleValSerProGlnAlaHisPhe 200
DB 894 GCTGGGACAAAGT 953
QY 201 PolyValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleArgCysGlnAspAla 220
DB 954 CTTAAGTCTCTCAATGCTTGAATATCAAGCGTCTTAAGTCAAGAAAGTGCAGATGCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleGlyValArgAspSer 240
DB 1014 TACCGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260
DB 1074 TCCAGGGGTGATTCGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1133
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysHis 280
DB 1134 TGGGAGATTTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1193

RESULT 15

US-09-991-073-308

/ Sequence 308, Application US/09991073

/ Patent No. US20020127576A1

/ GENERAL INFORMATION:

/ APPLICANT: Ashkenazi, Avi J.

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Botstein, David

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Eaton, Dan L.

/ APPLICANT: Ferrara, Napoleone

/ APPLICANT: Fong, Sherman

/ APPLICANT: Gerber, Hanspeter

/ APPLICANT: Gerritsen, Mary E.

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Grimaldi, J. Christopher

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Kijavir, Ivar J.

/ APPLICANT: Napier, Mary A.

/ APPLICANT: Pan, James

/ APPLICANT: Paoni, Nicholas F.

/ APPLICANT: Roy, Margaret Ann

/ APPLICANT: Stewart, Timothy A.

/ APPLICANT: Tamas, Daniel

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Williams, P. Mickey

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

/ FILE REFERENCE: P2730P1C15

/ CURRENT APPLICATION NUMBER: US/09/991,073

/ PRIOR FILING DATE: 2001-11-14

/ PRIOR APPLICATION NUMBER: 60/049787

/ PRIOR FILING DATE: 1997-06-16

/ PRIOR APPLICATION NUMBER: 60/062250

/ PRIOR FILING DATE: 1997-10-17

/ PRIOR APPLICATION NUMBER: 60/065186

/ PRIOR FILING DATE: 1997-11-12

/ PRIOR APPLICATION NUMBER: 60/065311

/ PRIOR FILING DATE: 1997-11-13

/ PRIOR APPLICATION NUMBER: 60/066770

/ PRIOR FILING DATE: 1997-11-24

/ PRIOR APPLICATION NUMBER: 60/075945

/ PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

| Pred. No.: | 4,45e-187 | Length: | 1570 |
|------------------------|-----------|---------------|------|
| Score: | 1608.00 | Matches: | 293 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x US-09-991-073-308 (1-1570)

QY 1 MetAlaThrAlaArgProProMetTTPValLeuCyAlaLeuIleThrAlaLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGAGGCTGCTGATCAAGCTTGCTT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 414 CTGGGGGTCAAGAGCATGTTCTGCCAACAATGATGTTCCGTGACCACTCTTAA 473
QY 41 ThrValProSerGlySerSerGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB 474 ACCGTGCCCTCTGGGAGCAACGAGACTGGAGCTGGGGCCGGGAGACGCCCGGTG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 534 GATGACAGACGACGCCGATCATCATGATCCGACTGGATATGCAACCAAGCCGTG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyCyGlyAlaValLeuValHisPro 100
DB 594 CAGGCCGCTGTGGCTTAAGGCCCAACAGCTCTATCGGGGGCGGTGGTGCATCCA 653
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyGlyValPheArgValArgLeuGlyHis 120
DB 654 CAGTGCTCTCTACGCGCCGCCCATGCAAGAAAGTTTCAGAGTCCGTCTCGGCCAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 714 TACTCCCTGTCAACAGTTTATGATTTGGGACAGATGTTCCAGGGGGTCAATCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLeuLeuAsn 160
DB 774 CCCCACTGGGTACCCACCTGGCACTCTAACGACCTCATGCTCAAACTGAAC 833
QY 161 ArgArgGlyLeuArgProThrIleAspValArgProIleAsnValSerSerHisCyAspSer 180
DB 834 AAGAAGATTGCTCCCACTAAAGATGCAAGACCATCAACGTCTCTCATTTGCTCTCT 893
QY 181 AlaGlyThrIleCysLeuValSerGlyTTPGlyThrThrIleSerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGCTTGGGTGTGGCTGGGGAACAACAAGAGCCCCCAAGTGCATTC 953
QY 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleArgCyGluAspAla 220
DB 954 CTTAAGTCTCTCAAGTCTTGAATATCAAGCTGCTAAGTCAAGAAAGTGCAGAGATGCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValArgAspSer 240
DB 1014 TACCCGAGACAGTAAATGACACCATGTTCTGGCCGGTGACAAACAGTAGAGACTCC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TGCAGGGGATGATTCGGGGGGCTGTGTCTGCAATGCTCCCTGCAAGGAACTGTGTCC 1133

QY 261 TTPGlyAspTyrProCyAlaIleArgProAsnArgProGlyValTyrThrAsnLeuCySlys 280
DB 1134 TGGGAGATTAACCTTGTGCCCCGCCCAACAGACCGGGGTCTCACGAACTTGCAAG 1193
QY 281 PheThrIleTTPIleGlnGluThrIleGlnAlaAsnSer 293
DB 1194 TTCAACCAAGTGATCCAGGAAACCATCCAGGCAACTCC 1232

Search completed: February 26, 2005, 22:39:46
Job time : 641 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:36:36 : Search time 3685 seconds
(without alignments)
3026.548 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608
Sequence: 1 MATRPRPMWVLCALITAL.....VYTLCKFTKMQIFTIQANS 293

Scoring table:

BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2_1/USPTO/US0936271/tunat_23022005_130127_27273/app_query.fasta_1.455
-DB=EST -QMT=FastGap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRB=US0936271.QCGN_1_1_3437_@tunat_23022005_130127_27273 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_g981:
9: gb_g982:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|-------------|
| 1 | 1248.5 | 77.6 | 960 | 2 | BE8899189 |
| 2 | 1158.5 | 72.0 | 1240 | 3 | AK003996 |
| 3 | 1143 | 71.1 | 1054 | 2 | BE747860 |
| 4 | 1103.5 | 68.6 | 997 | 6 | BY704245 |
| 5 | 1066.5 | 66.3 | 714 | 2 | BE898407 |
| 6 | 1041 | 64.7 | 729 | 2 | BE8842755 |
| 7 | 988.5 | 61.5 | 776 | 4 | BE680921 |
| 8 | 899 | 55.9 | 663 | 2 | BE742856 |
| 9 | 894 | 55.6 | 802 | 4 | BE682309 |

| | | | | | |
|----|-------|------|------|---|----------|
| 10 | 891 | 55.4 | 677 | 2 | BE388198 |
| 11 | 787 | 48.9 | 677 | 2 | BE898385 |
| 12 | 744 | 46.3 | 482 | 2 | AM820357 |
| 13 | 734 | 45.6 | 498 | 7 | W73168 |
| 14 | 734 | 45.6 | 533 | 4 | BM840509 |
| 15 | 734 | 45.6 | 577 | 4 | BM838406 |
| 16 | 734 | 45.6 | 590 | 4 | BM840511 |
| 17 | 730 | 45.4 | 526 | 4 | BM841697 |
| 18 | 724 | 45.0 | 405 | 4 | BM760479 |
| 19 | 722 | 44.9 | 544 | 5 | EX520121 |
| 20 | 720 | 44.8 | 666 | 4 | BM842155 |
| 21 | 719 | 44.7 | 607 | 7 | CV026544 |
| 22 | 713 | 44.3 | 953 | 2 | BE745465 |
| 23 | 705 | 43.8 | 523 | 4 | BM841293 |
| 24 | 698 | 43.4 | 512 | 4 | BM830263 |
| 25 | 690.5 | 42.9 | 557 | 4 | BM032286 |
| 26 | 680 | 42.3 | 809 | 4 | BG680075 |
| 27 | 671 | 41.7 | 765 | 9 | AY400701 |
| 28 | 661 | 41.1 | 765 | 9 | AY400702 |
| 29 | 655 | 40.7 | 457 | 1 | AI385433 |
| 30 | 650.5 | 40.5 | 1295 | 3 | AK009360 |
| 31 | 650.5 | 40.5 | 1295 | 3 | AK009720 |
| 32 | 641.5 | 39.9 | 783 | 9 | AY410898 |
| 33 | 634 | 39.4 | 765 | 9 | AY400703 |
| 34 | 632 | 39.3 | 1282 | 3 | BC015551 |
| 35 | 631.5 | 39.3 | 1269 | 3 | AK009659 |
| 36 | 627 | 39.0 | 783 | 9 | AY410900 |
| 37 | 627 | 39.0 | 880 | 6 | CB202840 |
| 38 | 621 | 38.6 | 898 | 6 | CB204935 |
| 39 | 620 | 38.6 | 945 | 3 | BU151248 |
| 40 | 619 | 38.5 | 852 | 6 | CB587168 |
| 41 | 618 | 38.4 | 1254 | 4 | AK087426 |
| 42 | 612.5 | 38.1 | 754 | 4 | BI653899 |
| 43 | 611.5 | 38.0 | 783 | 9 | AY410899 |
| 44 | 597 | 37.1 | 1447 | 3 | AK077406 |
| 45 | 597 | 37.1 | 1745 | 3 | AK029477 |

ALIGNMENTS

RESULT 1
BE8899189
LOCUS BE8899189 960 bp mRNA linear EST 29-SEP-2000
DEFINITION 601681989F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3952031 5', mRNA sequence.
ACCESSION BE8899189
VERSION BE8899189.1 GI:10366425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 960)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
AUTHORS Contact: Robert Straubeberg, Ph.D.
TITLE Email: cgapbs-r@mail.nih.gov
COMMENT Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/BLNT at: image.llnl.gov
Plate: LINC821 row: a column: 24
High quality sequence stop: 744.
Location/Qualifiers
1..960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3952031"
/tissue_type="adenocarcinoma cell line"

COMMENT

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, Riken Genomic Sciences Center (GSC), Riken Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGCGCGGCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3']. cDNA was cleaved with XhoI and SctI. Cloning sites, 5' end: SctI; 3' end: XhoI. Host: SOLR.

FEATURES

SOURCE

Location/Qualifiers

1..1240
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:1110030019"
/db_xref="taxon:10090"
/clone="1110030019"
/issue_type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
49..930
/note="unnamed protein product; putative
weakly similar to KALLIKREIN 5 PRECURSOR (EC 3.4.21.-)
(STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN
2) (KLC-12) [Homo sapiens] (SWISSPROT|Q9Y337, evidence:
FASTY, 69.7%ID, 99.6%length, match=879)"]
/codon_start=1
/protein_id="BAB23113.1"
/db_xref="GI:12834991"
/translocation="MARTGHMPKMAATLITLVLGVSEBVLADVSSCDNPGSTERS
/translocation="MARTGHMPKMAATLITLVLGVSEBVLADVSSCDNPGSTERS
GTRNDLSTDSKSDTRSDSSRIVNGSDOKDQAPQALILGPNLXGCAVLISPO
WLTLAHCRRKPRIRLGHSMSPVSSGOMQGISITHPGSHGHNDMLIKM
NRKIRDSHVKPVEIACDCATBGRVCSGMSGHTSSHNHPKYLQCLNTIVSEBIC
KNSYPGQIDKTFCADEBGRDSCQSDSGSPVVCNGKLQGLVSMGDPCKAQRNRPVY
TNLCFVKWIKDTMNSN"
1218..1223
/note="putative"
1240
/note="putative"

CDS

polYA_signal
polYA_site
polYA_site
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 4,79e-103 Length: 1240
Score: 1158.50 Matches: 205
Percent Similarity: 83.67% Conservative: 41
Best Local Similarity: 69.73% Mismatches: 45
Query Match: 72.05% Indels: 3
DB: 3 Gaps: 2

US-09-936-271b-14 (1-293) x AK003996 (1-1240)

QY 1 MetAlaThrAlaArgProTyrMetTyrValLeuCybAlaLeuLeuLeuLeu 20
Db 49 ATGGCGAGGACCGGACACCCCTGGAAATGGGCAATGCTACCCCTGATCACACCTGTT 108
QY 21 LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerAsn 40
Db 109 CTGGGGGCTTCAGAGCTGTTCTTGGGGGATTTTCTCTTGTGCAACCCCTTCGGA 168
QY 41 ThrValProSerGlySerAsnGlnAspLeuGly-----AlaGlyAlaGlyAlaAspAla 58
Db 169 ACCGAACTCTTGGGACCAACAGACCTGAGCAGGATTCGAAGTCTGGGGAGGACACC 228

QY 59 ArgSerAspAspSerSerSerArgIleIleLeuGlySerAspCysAspMetHisThrGln 78
Db 229 CGTTCA---GATGACGAGCTCTCAATTTGTAATGGGTCAGACGCAAAAGATGCACAG 285
QY 79 ProTyrGlnAlaAlaLeuLeuLeuAlaArgProAsnGlnLeuTyrCysGlyAlaValLeuVal 98
Db 286 CCAATGGCAGGGCCCTCTGCTTCTGGGGCCCAACACTGTACTGTGGGCGCTGTGATC 345
QY 99 HisProGlnTyrLeuLeuThrAlaAlaHisCysArgValLeuValPheArgValArgLeu 118
Db 346 AGCCCAACAGTGGCTGCTCAGCAGACACTGCAGAAAGCCAGTTCACAAATCCGCTG 405
QY 119 GlyHisTyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLe 138
Db 406 GGCACCATTCATGTCACCTGTTGATGATCTGGGCGACGAGATGTTCCAGGGAATCAAA 465
QY 139 SerIleProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLe 158
Db 466 TCCATCCCCCACCCTGGTTACTCCACCTGGGCACTCCCAATGACTCATGCTCATCAAA 525
QY 159 LeuAsnArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCys 178
Db 526 ATGAACAGAAATCCGTCAGCTCTCACTCAGTGAACCCGTCGAAATTGCTGTGACTGT 585
QY 179 ProSerAlaGlyTyrThrCysLeuValSerGlyTyrGlyTyrThrTyrSerProGlnVal 198
Db 586 GCCACCGAGGAGCTAGTGCATGTCGTCTGGGCGGCAACGACGACACCCACAT 645
QY 199 HisPheProValLeuGlnCysLeuAsnIleSerValLeuSerGlnTyrAspGlu 218
Db 646 AACTTCCGAAATCTCTCCAGTCCCTGATATTAAGTGTCTGCTAGAGAGAGTGA 705
QY 219 AspAlaTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspValGlyArg 238
Db 706 AACTCTCAACCAAGACAGATAGACAGACCATGTTCTGCCAGATGATGAAGGCGCAG 765
QY 239 AspSerCysGlnIleValSerSerGlyValProValValCysAsnIleSerLeuGlnIle 258
Db 766 GACTCTCTGCGGAGGTATTCGAGGCTCTGTCGATGCAATGCAATGAGGCGCTT 825
QY 259 ValSerTyrGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeu 278
Db 826 GTTCTCTGGGATATTTCTCTGCTGCTCAGCGGAACAGACAGAGTGTCTACACCACTG 885
QY 279 CysLeuPheThrTyrTyrIleGlnIleGlnIleGlnAlaAsn 292
Db 886 TGTGAGTTCGTTAAGTGAATTAAGACACCATGACTCCAACT 927

RESULT 3
BE747860
LOCUS
DEFINITION
601579720P1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928372 5',
mRNA sequence.
ACCESSION
BE747860
VERSION
BE747860.1 GI:10161852
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1054)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov.

Plate: LCM759 row: h column: 05
High quality sequence start: 16
High quality sequence stop: 632.
Location/Qualifiers

FEATURES

source

```
1..1054
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3928372"
/risue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_9"
/note="Organ: Ovary; Vector: pOTB7; Site: 1: XhoI; site: 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 1,27e-101 Length: 1054
Score: 1143.00 Matches: 230
Percent Similarity: 89.35% Conservative: 5
Best Local Similarity: 87.45% Mismatches: 22
Query Match: 71.08% Indels: 8
DB: 2 Gaps: 3
```

US-09-936-271B-14 (1-293) x BE747860 (1-1054)

```

QY 20 LeuLengIyValThrGluIuHlValLeuAlaAsnAspValSerCyAspHisProSer 39
DB 11 TTAAGTGGG-----GAGGGCAGTGTCTCCGCTTAAGATGTTCTCTGTACACCCCTCT 64
QY 40 AenThrValProSerGlySerAenGlnAenPleuGlyValAGlyValAGlyValAspAlaArg 59
DB 65 AACACCGTCCCTCTGGAGCAACACGAGCTGGAGCTGGGGCCGGGGAAGACGCCCG 124
QY 60 SerAspAspSerSerSerArgIleIleAenGlySerAspCysAspMetHisThrGlnPro 79
DB 125 TCGGATGACAGCAGCGCCGACATCATGATGATCCGACTGTCATGTGACACCCCGCG 184
QY 80 TTPGlnAlaAlaLeuLeuLeuArgProAenGlnLeuTyrCysGlyValAlaValLeuValHis 99
DB 185 TGGCAGGCGCGCTGTTGCTAAGGCCCAACACCTTACTGCGGGCGGCTTTGTCAT 244
QY 100 ProGlnTTPLeuLeuThraAlaAlaHisCysArgGlyValValPheArgValArgLeuGly 119
DB 245 CCACAGTGGCTGCTCAGCGCGCCGCTGACAGGAAGAAATTTCAGAGTCCGCTCGGC 304
QY 120 HisTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSer 139
DB 305 CACTACTCCCTGTGACACAGTTTATGATTTGGGCAGCAATGTTCCAGGGGCTC-AAATCC 363
QY 140 IleProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLeuLeu 159
DB 364 ATCCCCCAGCTGGCTACTCCACCTGGGCATCT-AAAGCACTCATGCTCATCAACAG 422
QY 160 AsnArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysPro 179
DB 423 AACAGAGAAGATTCGTCCCACTAAAGATGTCAAGCCATCAAGCTCTCCTCATTTGCC 482
QY 180 SerAlaGlyThrIlySerLeuValSerGly-TTPGlyThrThrIlySerProGlnValHis 199
DB 483 TCTGCTGGGACAAAGCTTGTGCTGGCTGGGGGACAAACAAAGCCCCCAACTGGA 542
QY 199 sPheProIlyValIleGlnCysLeuAsnIleSerValIleSerGlnIlyCysGlnAs 219
DB 543 CTTCCTTAAGTCATCACTGCTGTGAATTCACCGCTGTAAGTCCGAAACAGGTGCAGGA 602
QY 219 palATyrProArgGlnIleAspAspThrMetPheCysAla-GlyAspIlyAlaGlyArg 239
```

```

DB 603 TGTCTACCCGAGACAGATAGATGACACCATGACTGCGCCCGGTGACAAAGAGGTAGA 662
QY 239 sPseSerCyGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeu 259
DB 663 GACTCATGCAGAGTATCTTGGGGGGCTTGTGTGTGAATGCTCTCGAGGAAATCGTG 722
QY 259 aIseTTPGlyValPlyTyrProCysAlaArgProAsnArgProGlyValTyrThrAenLeu 279
DB 723 TC---TGGGGAGATTACCTTGTGCA---GGCATTAGACCGGGTGTGTACAGAGAACTTC 776
QY 279 yAlAs 280
DB 777 GGAAG 781

RESULT 4
BY704245
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BY704245 997 bp mRNA linear EST 16-DEC-2002
BY704245 RIKEN full-length enriched, 18-day embryo whole body Mus
musculus cDNA clone 1110030019 5', mRNA sequence.
BY704245
BY704245.1 GI:27115369
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 997)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamataka, T.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kapustin, A., Marcande, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Brd, D., Brusic, V.,
Chochia, C., Corbani, L. E., Cousins, S., Della, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gariboldi, M., Glasi, C., Godzik, A., Gough, J., Gimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Matsuda, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. V., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Veiersted, R., Wagner, L., Wahlestedt, C., Wang, Y., Watnabe, Y.,
Wells, C., Wilming, L. G., Wyszewski, B. A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Waki, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, T., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
JOURNAL
MEDLINE
PUBMED
COMMENT
```

CONTACT: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, K., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kawai, Y., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Source

Location/Qualifiers

```
1..997
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1110030019"
/tissue_type="whole body"
/dev_stage="18-day embryo"
/clone_1lb="RIKEN full-length enriched, 18-day embryo whole body"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 9.02e-98 Length: 997
Score: 1103.50 Matches: 200
Percent Similarity: 82.99% Conservative: 39
Best Local Similarity: 69.44% Mismatches: 46
Query Match: 68.63% Indels: 4
DB: 6 Gaps: 2
```

US-09-936-271b-14 (1-293) x BY704245 (1-997)

```
QY 1 MecalathrilaahgProptTmPctTValLeuCybAlaLeuIethrAlaLeu 20
DB 47 ATGGCGAGGACCGGACACCCCTGGAATGGCAATGGCTACCTGATCACAACCTGCTT 106
QY 21 LeuGlyValThrgIuHsValLeuAlaAsnAmpValSerCybAAPHISProSerAsn 40
DB 107 CTGGGGGCTCAGAGCCTGTTCTTGGGGATGTTCTCTTGAGAACCCCTCTGGA 166
QY 41 ThrValProSerGlySerAsnGlnAmpLeuGly-----AlaGlyAlaGlyGluAspAla 58
DB 167 ACGGAACCCCTGCGACCAAGACCTCAGACGAGTTCAAGTCTGGGGAGAGACACC 226
QY 59 ArgSerAmpAmpSerSerSerArgIleIleAenGlySerAmpCybAmpMetHisThGln 78
DB 227 CGTTCA---GATACACACTCTCGAATTGTGAATGGCTCAGACTCCAAAAGATGCAG 283
QY 79 ProTtPglAlaAlaLeuLeuValArgProAenGlnLeuTyrCybGlyAlaValLeuVal 98
DB 284 CCAATGGGAGGGGGCGCTCTCTGGGGCCCAAGAGCTGATCTGGGGCTGCTGATC 343
QY 99 HisProGlnTtPLeuLeuThrAlaAlaHisCybArgGlyValValPheArgValArgLeu 118
DB 344 AGCCACAGTGGCTGCTCAGCAGACACCTGCAAGAAAGCCAGTTCAGAAATCGCTG 403
QY 119 GlyHisTyrSerLeuSerProValTyrGlySerGlyGlnMetPheGlnGlyValLys 138
DB 404 GGCACCATTCACATTCACCTGCTATGAGTCTGGGCGAGAGATGTTCCAGGGATCAAA 463
QY 139 SerIleProHisProGlyTyrSerHisProGlyHisSerAmpAmpLeuMetLeuIleLys 158
```

```
DB 464 TCATATCCCAACCCCGGTTACTCCACCTCGGCGACATGCATGACCTCATGCTATCAA 523
QY 159 LeuAmpAmpArgIleArgProThrLysAmpValArgProIleAenValSerSerHisCys 178
DB 524 ATGAACGAAATAATCCCTGACTCTCAGTCAAGTAAAGCCGTGAAATTGCTTGACCTG 583
QY 179 ProSerIleArgLysLysCybLeuValSerGlyTtPglLysThrLysSerProGlnVal 198
DB 584 GCCACCGAGGAGACTAGTGCATGAGTCTGGGCTGGGGAGCAAGACAGCAGCACAAT 643
QY 199 HisPheProLysValLeuGlnCysLeuAenIleSerValLeuSerGlnTyrArgCysGlu 218
DB 644 AACTCCCGAAGAGCTCCAGTGTGCTGAATTAATGCTGCTGAGTGAAGAGGTGTA 703
QY 219 AspAlaTyrProArgGlnIleAmpAmpThrMetPheCysAlaGlyAmpLysAlaGlyArg 238
DB 704 AACTCTTACCCAGAGACAGATGACAAAGACCATGTTCTGCCAGGTGATGAAGGGCAGG 763
QY 239 AspSerCybGlnGlyAmpSerGlyGlyProValValCybAenGlySerLeuGlnGlyLeu 258
DB 764 GACTCTCGCAGGAGGTGATTCGGAGT-CTGTGCTGCAATGGCAATGACAGGGGCTT 822
QY 259 ValSerTtPglLysAmpTyrProCybAlaArgProAmpAmpProGlyValTyrThrAsnLeu 278
DB 823 GTGTCTGGGGGTATTTCCCTGCTGCTCAGCGAAGACAGACAGTGTCTACACCCACTG 882
QY 279 CysLysPheThrLysTtPglLeu 286
DB 883 TGTGAGTTCGTGAATGATTTAG 906
```

RESULT 5

BE898407

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

```
1..714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951613"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 9"
/note="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit
```

ORIGIN (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2,44e-94 | Length: | 714 |
| Score: | 1066.50 | Matches: | 214 |
| Percent Similarity: | 97.72% | Conservative: | 0 |
| Best Local Similarity: | 97.72% | Mismatches: | 4 |
| Query Match: | 66.32% | Indels: | 5 |
| DB: | 2 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x BE898407 (1-714)

```
QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
DB 64 ATGGCTACAGCAAGACCCCTGGAGTGGTGTCTGTGCTGTATCAAGCTTGCTT 123
QY 21 LeuGlyValThrGluHisValLeu-AlaAsnAsnAparValSerCyAspHisProSerAs 40
DB 124 CTGGGGGTCAAGAGCATGTCTCCGGCCAATGATGTTTCTGTGACCACTCTTA 183
QY 40 nThrValProSerGlySerAsnGlnAparLeuGlyAlaGlyAlaGlyAlaArgSe 60
DB 184 CACCGTCCCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAAAGACCGCGT 243
QY 60 rAspAspSerSerSerArgIleIleAsnGlySerAparCyAspMetHisThrGlnProTr 80
DB 244 GGATGACACAGAGAGCGCATCATCATATGATTCGATCGATGATGACACCCAGCCG 303
QY 80 pGlnAlaAlaLeuLeuAparProAsnGlnLeuTyrcGlyAlaValLeuValHisPr 100
DB 304 GGAGCGCGGCTGTCTTAAGGCCCAACGAGCTCTAGCGGGCGGTGTGGTGCATCC 363
QY 100 oGlnTrpLeuLeuThrAlaAlaHisCyAsnGlySerValPheArgValArgLeuGlyHis 120
DB 364 ACGAGTGGCTCAGCGCGCCGACCTGCGAGAAAGATTTCAGATCCGCTCGGCCA 423
QY 120 sTyrcSerLeuSerProValTyrcGluSerGlyGlnGlnMetPheGlnGlyValLysSerI 140
DB 424 CTACTCCCTGTACACGATTATGAACTCGGCGAGCGAGATTTCCAGGGGTCC-AAATCCAT 482
QY 140 eProHisPProGlyTyrcSerHisPProGlyHisSerAAsnAparLeuMetLeuIleLysLeuAs 160
DB 483 CCCCCACCTGGCTACTCCACCCCTGG-CACCTTAACGACCTCATCTCATCAAACTGAA 541
QY 160 nArgArgIleArgProThrLysAparValArgProIleAsnValSerSerHisGlySerProSe 180
DB 542 CAGAGAAATTCGCCCACTAAAGATGTCAGACCCCATCAACGTCTCTCATTTGTCCTTC 601
QY 180 rAlaGlyThrLysCyAlaLeuValSerGlyTyrcGlyThrThrLysSerProGlnValHisPh 200
DB 602 TCGTGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
QY 200 eProLysValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnLysArgCyAsnGlu 218
DB 660 CCTTAAGGTCTCCAGTGTCTTAATATCATGCGGTAAAGTCAAGAAAGTGGCGAG 714
```

RESULT 6
LOCUS BM842755 729 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0120216 S12SNU216 Homo sapiens cDNA clone S12SNU216-66-100 5',
mRNA sequence.

ACCESSION BM842755
VERSION BM842755
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 729)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krbb.re.kr
Plate: 66 row: D column: 10
High quality sequence stop: 729.
Location/Qualifiers

FEATURES

source

1..729

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="S12SNU216-66-D10"  
/sex="F"  
/tissue_type="Lymph node"  
/cell_type="Epithelial"  
/cell_line="SNJ-216"  
/lab_host="Top10P"  
/clone_lib="S12SNU216"  
/note="Organ: Stomach; Vector: pcNS; Site: 1: EcoRI;  
Site: 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10P by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."
```

ORIGIN

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 8.1e-92 | Length: | 729 |
| Score: | 1041.00 | Matches: | 191 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 64.74% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x BM842755 (1-729)

```
QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
DB 156 ATGGCTACAGCAAGACCCCTGGAGTGGTGTCTGTGCTGTATCAAGCTTGCTT 215
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAparValSerCyAspHisProSerAsn 40
DB 216 CTGGGGGTCAAGAGCATGTCTCCGCAATGATGTTTCTGTGACCACTCTTAAC 275
QY 41 ThrValProSerGlySerAsnGlnAparLeuGlyAlaGlyAlaGlyAlaArgSer 60
DB 276 ACCGTGCCCTTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAAAGACCGCGTGG 335
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAparCyAspMetHisThrGlnProTrp 80
DB 336 GATGACAGAGAGAGCGCATCATCATATGATCGACTGGGATATGACACCAAGCGGTGG 395
QY 81 GlnAlaAlaLeuLeuAparProAsnGlnLeuTyrcGlyAlaValLeuValHisPro 100
DB 396 CAGGCGCGCTTTGCTAAGGCCCAACAGCTCTACTGGGGGGCGGTGTGGTGCATCCA 455
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAsnGlySerValPheArgValArgLeuGlyHis 120
```

Db 456 CAGTGGCTGCTCAGCGCCGCTGAGGAGAAAGTTTTCAGATCCGCTCGGCGAC 515
 Qy 121 TyrsertleuserProvalTYrGIuserGIyGInGInmetPheGInGlyValIySerIle 140
 Db 516 TACTCCCTGTCACAGGATTTGATGATCTGGGACAGAGATGTTCCAGGGGGTCAATTCATC 575
 Qy 141 ProHIsProGIyTYrSerHIsProGIyHIsSerHIsPLeuMetLeuIleIyLeuAsn 160
 Db 576 CCCGACCTGCTGCTCTCCACCTGCGCACTTACACCTCATGCTCATCAACTGAC 635
 Qy 161 ArgArgIleArgProThryAspValArgProIleAnValSerSerHIsCyPProSer 180
 Db 636 AGAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATTTGCTCTCT 695
 Qy 181 AlAGIyThryCyLeuValSerGIyTYrGIy 191
 Db 696 GCTGGGACAAAGCTTGCTGTCTGCTGGG 728
 RESULT 7
 LOCUS BG680921 776 bp mRNA linear EST 01-MAY-2001
 DEFINITION 602648718F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753469 5',
 mRNA sequence.
 ACCESSION BG680921
 VERSION BG680921.1 GI:13912305
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 776)
 NIH-MGC http://mgs.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM10613 row: C column: 06
 High quality sequence stop: 770.
 Location/Qualifiers
 1..776
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4753469"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NCI CGAP Skn4"
 /notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 ORIGIN
 Alignment Scores:
 Prod. No.: 13e-86 Length: 776
 Score: 988.50 Matches: 197
 Percent Similarity: 93.87% Conservative: 2
 Best Local Similarity: 92.92% Mismatches: 7
 Query Match: 61.47% Indels: 6
 DB: 4 Gaps: 1
 US-09-936-271b-14 (1-293) x BG680921 (1-776)
 Qy 1 MetAlaThraIaaArgProTrpMetTrpValIeuCysAlaIeuIleThraIaIeuIeu 20
 Db 151 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTCTGTGATCACAGCCTTCTT 210

Qy 21 LeuGIyValThrGIuHIsValIeuAlaAsnAspValIserCysAspHIsProSerAsn 40
 Db 211 CTGGGGGTCAAGAGCATGTTCTCGCAACAAAGATGTTCTGTGACACCCCTTAC 270
 Qy 41 ThrValProSerGIySerAenGIaAspLeuGIyAlaGIyValIaGIyIuAspAlaArgSer 60
 Db 271 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGCCCGGGAAGACCCGCTG 330
 Qy 61 AspAspSerSerArgIleIleangIySerAspCysAspMetHIsThrGIaProTrp 80
 Db 331 GATGACAGCAGACCGGCATCATGATGATCGAGCTGATATGACACCCGCTGG 390
 Qy 81 GIaAlaAlaIeuIeuIeuAArgProAnGIuIeuTYrCyGIyValIaValIeuValHIsPro 100
 Db 391 CAGGCGGCTGTGTGCTAAGCCCAACAGCTTCTTCCGGGCGGTGTGGATCCA 450
 Qy 101 GIaTrpIeuIeuThraIaAlaHIsCyArgIyIyValIaPheArgValATrLeuGIyHIs 120
 Db 451 CAGTGGCTGCTCAGCGCGCCGCTGAGGAAAGATTTTCAGAGTCCGCTCGGCGAC 510
 Qy 121 TyrsertleuserProvalTYrGIuserGIyGInGInmetPheGInGlyValIySerIle 140
 Db 511 TACTCCCTGTCACAGATTTGATGATCTGGGACAGATGTTCCAGGGGGTCAATTCATC 570
 Qy 141 ProHIsProGIyTYrSerHIs-ProGIy-HIsSerAsp-LeuMetLeuIleIyLeu 159
 Db 571 CCCCACTGCTGCTACCGCCGCTGAGCAGCTTAACGAGCTCATGATCAACTG 630
 Qy 160 -AaaArgArgIleArgProThryAspVal-ArgProIleAnValSerSerHIsCysP 179
 Db 631 AAACAGAGAATTCGTCCCACTAAAGATGTCAGACCCATCAAGTCTCTCATGTC 690
 Qy 179 roSerAlaGIyThryCyLeuValSerGIyTYrGIyThryThrySerProGIaValH 199
 Db 691 CCTCTGCTGGAGCAAGATGCTGTGTGCTGGGAGACAGACCAAGACCCCAAGTGC 750
 Qy 199 IasPeproIyValIeuGIuIeuVal 207
 Db 751 ACTTCGCTAGG--TCTCAGTCTG 773
 RESULT 8
 LOCUS BE742856 663 bp mRNA linear EST 15-SEP-2000
 DEFINITION 601574513F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3835490 5',
 mRNA sequence.
 ACCESSION BE742856
 VERSION BE742856.1 GI:10156848
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 663)
 NIH-MGC http://mgs.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: LLM517 row: J column: 03
 High quality sequence stop: 657.
 Location/Qualifiers
 1..663
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3835490"

/cissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH-MGC 9"
/notes="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 6,6e-78 | Length: | 663 |
| Score: | 899.00 | Matches: | 193 |
| Percent Similarity: | 93.72% | Conservative: | 1 |
| Best Local Similarity: | 93.24% | Mismatches: | 10 |
| Query Match: | 55.91% | Indels: | 8 |
| DB: | 2 | Gaps: | 1 |

US-09-936-271B-14 (1-293) x BG682309 (1-663)

QY 1 MetAlaThraAlaArgProPrtMetTrpValLeuCysAlaLeuIleThraAlaLeu 20
DB 54 ATGGCTACAGCAAGACCCCTGGATGGGTCTCTGCTCTGATCACAGCCTTGCT 113
QY 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCysAspHisProSerAn 40
DB 114 CTGGGGGTACAGAGATGTTCTGGCCAAAGAGATTTCTGTATACCACTCTTATAC 173
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValAlaGlyValAspValSer 60
DB 174 ACCGTCTCTCTGGAGCAACCGACCTGGAGCTGGGGCCGGGAACGCCCTGGT 233
QY 61 AspAspSerSerGlyIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 234 GATGACAGACAGCGCGCATCATGATCGACTGCGATATGACACCCAGCGTGG 293
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyValValLeuValHisPro 100
DB 294 CAGGCGCGCTGTGTCTAAAGCCCAACAGCTTACTGGGGGGGTGTGGTGCATCCA 353
QY 101 GlnTrpLeuLeuThraAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
DB 354 CAGTGGCTCTCAAGCGCCG-CACTGCAAGAAAGATTTCAGATCCCTCTCGGCCAC 412
QY 121 TyrBerLeuSerProValIleArgIleSerGlyGlnIleMetPheGlnGlyValIleSerTle 140
DB 413 TACTCCCTGTCAACCACTTATGATCTGGGACAGATGTTCCAGGGGTC-AAATCCATC 471
QY 141 ProHisProGlyTyrSerHisPProGlyHisSerAspAspLeuMetLeuIleLys 160
DB 472 CCCCACTGGGTATCCCA--CCTGGCCACTTAAAGACTCATGCTCACTCAAACTTGA 528
QY 160 nATGATGileArgProThrIleAspValArgProIleAsnValSerSerHisCysProse 180
DB 529 CAGAAATTCGTCCTCACTAAAGATGTCAGACATC-AACGTCTCTCTCATATTGATCC 587
QY 180 rAlaGlyThrIleCysLeuVal-SerGlyTrpGlyThrThrIleSerProGlnValHisP 200
DB 588 TGCTGGAGCAAGATGCTGTGTGTCTGGCTGGGGGACAC-AGAGGCCCAAGTGCAT 646
QY 200 heProIleValLeuGln 205
DB 647 C-CTTAAGGCTTGTAG 662

RESULT 9 802 bp mRNA linear EST 01-MAY-2001
LOCUS BG682309
DEFINITION 602629626f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',
ACCESSION BG682309 mRNA sequence.

VERSION BG682309.1 GI:13913706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL 1 (bases 1 to 802)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contract: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10615 row: d column: 08
High quality sequence stop: 802.

FEATURES

source

1..802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754263"
/cissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2.69e-77 | Length: | 802 |
| Score: | 894.00 | Matches: | 161 |
| Percent Similarity: | 99.38% | Conservative: | 0 |
| Best Local Similarity: | 99.38% | Mismatches: | 1 |
| Query Match: | 55.60% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-936-271B-14 (1-293) x BG682309 (1-802)

QY 132 GlnMetPheGlnGlyValIleSerIleProHisProGlyTyrSerHisProGlyHisSer 151
DB 3 CAGATGTTCCAGGGGTCAAAATCCATCCCACTGGCTACTCCACCTGGCCACTCT 62
QY 152 AsnAspLeuMetLeuIleLysLeuAsnArgArgIleArgProThrIleAspValArgPro 171
DB 63 AACGACCTCATGCTCATCAACTGACAGAAAGATTCCTCCACTAAAGATGTACAGCC 122
QY 172 IleAsnValSerSerHisCysProSerAlaGlyThrIleCysValLeuValSerGlyTrpGly 191
DB 123 ATCAAGTCTCTCTCATATGTCCTCTGTGGAGCAAGTGTGTGGTCTGGCTGGGG 182
QY 192 ThrThrIleSerProGlnValHisPheProIleValLeuGlnCysLeuAsnIleSerVal 211
DB 183 ACAACCAAGAGCCCCCAAGTGCATTCCTTAAGGTCTCCAGTGTGTGAATATCAGCCGG 242
QY 212 LeuSerGlnLysArgCysGlnAspAlaTyrProArgGlnIleAspAspThrMetPheCys 231
DB 243 CTAACTCAAAAGGCGAGAGTGTACCCGAGACAGATGATGACACATGTTCTGC 302
QY 232 AlaGlyAspLysAlaGlyValArgAspSerCysGlnGlyAspSerGlyValProValValCys 251
DB 303 GCGGTGACAAAGCAGGTAGAGATCTCTCCAGGGGATTCGGGGGGCTGTGGTCTGC 362
QY 252 AangIySerLeuGlnGlyLeuValSerTrpGlyAspTyrProCysAlaArgProAsnArg 271
DB 363 AATGGCTCCCTCAGGAGACTCGTGTCTGGGAGATTACCTTGTGCCCCGCGCCAAAGAG 422

QY 272 ProGlyValThrAspLeuCyLeuPheThrTrpIleGlnIleGlnAla 291
|||||
Db 423 CCGGGTCTTCAACGAACCTCTGCAAGTTCCAAAGTGATCCAGAAACATCCAGGCC 482
|||||
QY 292 AsnSer 293
|||||
Db 483 AAGTCC 488
|||||
RESULT 10
BE388198 677 bp mRNA linear EST 21-JUL-2000
LOCUS 60126704F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621682 5',
DEFINITION mRNA sequence.
ACCESSION BE388198.1 GI:9333563
VERSION BE388198.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 677)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Straube, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM295 row: m column: 11
High quality sequence stop: 618.
Location/Qualifiers
1. 677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3621682"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ALIGNMENT SCORES:
Pred. No.: 4.17e-77 Length: 677
Score: 891.00 Matches: 176
Percent Similarity: 92.15% Conservative: 0
Best Local Similarity: 92.15% Mismatches: 10
Query Match: 55.41% Indels: 7
DB: 2 Gaps: 1
US-09-936-271b-14 (1-293) x BE388198 (1-677)

QY 1 MetaLathrAlaArgProTyrMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
|||||
Db 101 ATGGCTACAGCAAGACCCCTGATGCTGCTCTGTGATCAGACGCTTCTT 160
|||||
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
|||||
Db 161 CTGGGGGTCAAGAGCATGTTCTGCAACAAATGATGTTCTTCTGACCAACCCCTTAAC 220
|||||
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60
|||||

Db 221 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAGAGCCGGCTCG 280
|||||
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCyAspMetHisTrpGlnProTyr 80
|||||
Db 281 GATGACAGCAGACCGGCATCATCATGATGCCAGTATGACACACCGCCGTGG 340
|||||
QY 81 GlnAlaAlaLeuLeuAsnGProAsnGlnLeuTyrCyGlyAlaValIleValHisPro 100
|||||
Db 341 CAGGCCGCGCTGTGCTTAAGCCCAACACACTTACTGCGGGGCGGTGGTCATCA 400
|||||
QY 101 GlnTrpLeuThrAlaAlaHisCyAspGlyLeuValPheArgValArgLeuGlyHis 120
|||||
Db 401 CAGTGGCTGCTCAAGCGCCGCTCAAGAGAAAGATTTCAGAGTCCGCTCGGCAC 460
|||||
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValSerIle 140
|||||
Db 461 TACTCCCTGCACAGATTATGATCTGGGACAGATGTTCCAGGGGGTCAATTCCAT 520
|||||
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValAsn 160
|||||
Db 521 CCCCACCTGG-7ACTCCACCTGGCCACTTACGACCTTACGATCAACTAAC 579
|||||
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyProSer 180
|||||
Db 580 AGAAGATTGCTCCATTAAAGATGTCAGACCATCAACGCTCTCTATGCT-CTCTG 638
|||||
QY 181 AlaGlyThrIleCyLeuValSerGlyTyrGly 191
|||||
Db 639 CTGGGA-----AAAGCTGGGT 656
|||||
RESULT 11
BE898385 677 bp mRNA linear EST 29-SEP-2000
LOCUS 601681219F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951365 5',
DEFINITION mRNA sequence.
ACCESSION BE898385.1 GI:10364809
VERSION BE898385.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 677)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Straube, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LICM619 row: f column: 06
High quality sequence stop: 677.
Location/Qualifiers
1. 677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951365"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

ORIGIN (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 7,116-67 | Length: | 677 |
| Score: | 787.00 | Matches: | 143 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 48.94% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-936-271b-14 (1-293) x BE898385 (1-677)

```
QY 151 SerAsnAspLeuMetLeuIleValLeuAsnArgArgIleArgProThrIleAspValArg 170
    |||||
DB 3 TCCTAAGACCTCATGCTCATCAACAGAGAGAAATTCCTCCATAAGATGTCAGA 62
    |||||
QY 171 ProIleAsnValSerSerHisCysProSerAlaGlyThrIleCysLeuValSerGlyTTP 190
    |||||
DB 63 CCCATCAAGCTCTCTCTCATTTGCTCCCTGCTGGAGCAAAAGCTTGGTCTGGCTGG 122
    |||||
QY 191 GlyThrThrIleSerProGlnValHisPheProIleValLeuGlnCysLeuAsnIleSer 210
    |||||
DB 123 GGGACACCAAGAGCCCAAGTCACTTCCCTCAAGGTCCTCAAGTCTGAATATCAGC 182
    |||||
QY 211 ValLeuSerGlnIleValArgCysGluAspAlaIleProArgGlnIleAspAspThrMetPhe 230
    |||||
DB 183 GTGCTAAGTCAAGAAAGAGTCCAGAGATGCTTACCCAGACAGATGATGACACCATGTC 242
    |||||
QY 231 CysAlaGlyAspIleValIleGlyArgAspSerCysGlnGlyAspSerGlyGlyProValIle 250
    |||||
DB 243 TGGCCCGGTGACAAACAGATGAGACTCCTGTCAGAGGTGATTTCTGGGGGCTGTGGTGC 302
    |||||
QY 251 CysAsnGlySerLeuGlnGlyLeuValSerTgplAspIleProCysAlaAspProAsn 270
    |||||
DB 303 TGGAAATGGCTCCCTGACGGAGCTGCTGCTGGAGAGATTAACCTTGTGCCCCGCCAC 362
    |||||
QY 271 ArgProGlyValIleThrAsnLeuCysIlePheThrIleGlnIleGlnIleGln 290
    |||||
DB 363 AGACCGGGGTCTCAACAGAACCTCTGCAAGTTACCAAGTGCATCCAGAAACATCCAG 422
    |||||
QY 291 AlaAsnSer 293
    |||||
DB 423 GCCAACTCC 431
    |||||
```

RESULT 12
AM820357/c 482 bp mRNA linear EST 17-MAY-2000

LOCUS AM820357
DEFINITION QV2-ST0296-100400-130-d07 ST0296 Homo sapiens cDNA, mRNA sequence.

AM820357
VERSION AM820357.1 GI:7913351

KEYWORDS EST.

SOURCE Homo sapiens (human)

REFERENCE
AUTHORS Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 482)
Dias Neri, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, R.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.J.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV2-ST0296-100

400-130-d07&t3=2000-04-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 62

High quality sequence stop: 482.

FEATURES

source

```
1..482
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /dev_stage="Adult"
    /clone_lib="ST0296"
    /note="Organ: stomach; Vector: puc18; Site 1: Smal;
    Site 2: SmaI; A mini-library was made by cloning products
    derived from ORESTES PCR (U.S. Letters Patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the pUC 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under
    low stringency conditions."
```

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 7,416-63 | Length: | 482 |
| Score: | 744.00 | Matches: | 139 |
| Percent Similarity: | 97.90% | Conservative: | 1 |
| Best Local Similarity: | 97.20% | Mismatches: | 2 |
| Query Match: | 46.27% | Indels: | 1 |
| DB: | 2 | Gaps: | 0 |

US-09-936-271b-14 (1-293) x AM820357 (1-482)

```
QY 152 AsnAspLeuMetLeuIleValLeuAsnArgArgIleArgProThrIleAspValArgPro 171
    |||||
DB 465 AACGACCTCATGCTCATCAACAGAGAAATTCCTCCATAAGATGTCAGAGCC 406
    |||||
QY 172 IleAsnValSerSerHisCysProSerAlaGlyThrIleCysLeuValSerGlyTTPgly 191
    |||||
DB 405 ATCAAGCTCTCTCAATGCTCCTGCTGAGACAAAGTGTGTGCTGGCTGGGG 346
    |||||
QY 192 ThrThrIleSerProGlnValHisPheProIleValLeuGlnCysLeuAsnIleSerVal 211
    |||||
DB 345 ACACCAAGAGCCCCCAAGTGCATCTCCCTTAAGTCTTCAAGTCTTGAATATCAGCGT 286
    |||||
QY 212 LeuSerGlnIleValArgCysGluAspAlaIleProArgGlnIleAspAspThrMetPheCys 231
    |||||
DB 285 CTAAAGTCAAGAAAGGTGCGAGAGATGCTTACCCGAGACAGATGATGACACCATGTTCTGC 226
    |||||
QY 232 AlaGlyAspIleValIleGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCys 251
    |||||
DB 225 GCCGGTGAACAAAGCGTGAAGAGATCTTCCACAGGTGATTCGGGGGCTGTGTCTGC 166
    |||||
QY 252 AangIleSerLeuGlnGlyLeuValSerTgplAspIleProCysAlaAspProAsnArg 271
    |||||
DB 165 AATGGCTCTCTCAAGGAGATCTGCTGCTGGAGATTAACCTTGTGCCCCGCCCAACAGA 106
    |||||
QY 272 ProGlyValIleThrAsnLeuCysIlePhe-ThrIleSTPILeglnIleGlnIle 291
    |||||
DB 105 CTGGGGTGTACACGAACTCCGCGAGTTCAACAAAGTGAATCCAGGAACATCCAGGC 46
    |||||
QY 291 AlaAsnSer 293
    |||||
DB 45 CAATCTCC 39
    |||||
```

RESULT 13

LOCUS W73168

DEFINITION

W73168 498 bp mRNA linear EST 16-OCT-1996
zds5e11.r1 Soares fetal heart_NbH19w Homo sapiens cDNA clone
IMAGE:344588 5' similar to SW:TR2Y_CNFPA P06872 TRYP5INOGN,

Alignment Scores:

Pred. No.: 8,25e-62 Length: 533
 Score: 734.00 Matches: 132
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.65% Indels: 0
 DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x BM840509 (1-533)

QY 162 Arg11LeuArgProThrLysAspValArgProIleAsnValSerSerHisCysProSerAla 181
 |||||
 DB 3 AGAATTCGTCCTCCCTAAAGATGTCAGACCCATCAAGCTCCCTCATTTGCTCTGCT 62
 QY 182 GlyThrLysCysLeuValSerGlyTTPGlyThrThrLysSerProGlnValHisPhePro 201
 |||||
 DB 63 GGGACAAAGTCTGGTGGTGGTGGTGGGACCAAGAGCCCAAGTGCATTCCT 122
 QY 202 LysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlnAspAlaTyr 221
 |||||
 DB 123 AAGTCTCTCCAGTCTTGAATATCAAGCTGCTAAGTCAGAAAAGTGCAGAGATGCTTAC 182
 QY 222 ProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSerCys 241
 |||||
 DB 183 CCGAGACAGATGATGACACACATGTTCTGCGCGGTGACAAACAGTAGAGACTCTGC 242
 QY 242 GlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSerTTP 261
 |||||
 DB 243 CAGGGGATCTGGGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 302
 QY 262 GlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLysPhe 281
 |||||
 DB 303 GGAAGATTACCTTGTGCTCCCGGCCCAACAGCCGGGTGTACAGAACCTTGCAAGTTC 362
 QY 282 ThrLysTTP11LeGlnGluThr11LeGln11aAsnSer 293
 |||||
 DB 363 ACCAAGTGAATCCAGAAACCATCCAGGCCAAGCTCC 398
 RESULT 15
 BM838406 577 bp mRNA linear EST 06-MAR-2002
 LOCUS K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5',
 DEFINITION mRNA sequence.
 ACCESSION BM838406
 VERSION BM838406.1 GI:19194815
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 577)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,U.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: A column: 03
 High quality sequence stop: 577.
 Location/Qualifiers
 1..577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SNU216-36-A03"
 /sex="P"

ORIGIN

/tissue_type="lymph node"
 /cell_type="Epithelial"
 /cell_line="SNJ-216"
 /lab_host="T0p10P"
 /clone_id="S12SNU216"
 /note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deacapped
 with tabacco acid pyrophosphatase (TAP). The deacapped
 intact mRNA was ligated with DNA-RNA linker including Scori
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli T0p10P by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

Alignment Scores:

Pred. No.: 9,25e-62 Length: 577
 Score: 734.00 Matches: 132
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.65% Indels: 0
 DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x BM838406 (1-577)

QY 162 Arg11LeuArgProThrLysAspValArgProIleAsnValSerSerHisCysProSerAla 181
 |||||
 DB 3 AGAATTCGTCCTCCCTAAAGATGTCAGACCCATCAAGCTCTCTCATTTGCTCTGCT 62
 QY 182 GlyThrLysCysLeuValSerGlyTTPGlyThrThrLysSerProGlnValHisPhePro 201
 |||||
 DB 63 GGGACAAAGTCTGGTGGTGGTGGTGGGACCAAGAGCCCAAGTGCATTCCT 122
 QY 202 LysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlnAspAlaTyr 221
 |||||
 DB 123 AAGTCTCTCCAGTCTTGAATATCAAGCTGCTAAGTCAGAAAAGTGCAGAGATGCTTAC 182
 QY 222 ProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSerCys 241
 |||||
 DB 183 CCGAGACAGATGATGACACCATGTTCTGCGCGGTGACAAAGCAGTAGAGACTCTGC 242
 QY 242 GlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSerTTP 261
 |||||
 DB 243 CAGGGTGAATTCGGGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 302
 QY 262 GlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLysPhe 281
 |||||
 DB 303 GGAAGATTACCTTGTGCTCCCGGCCCAACAGCCGGGTGTACAGAACCTTGCAAGTTC 362
 QY 282 ThrLysTTP11LeGlnGluThr11LeGln11aAsnSer 293
 |||||
 DB 363 ACCAAGTGAATCCAGAAACCATCCAGGCCAAGCTCC 398

Search completed: February 26, 2005, 22:29:06
 Job time : 3693 secs